

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
3 April 2003 (03.04.2003)

PCT

(10) International Publication Number
WO 03/027274 A2

(51) International Patent Classification⁷: C12N 9/90,
C07K 14/21, 19/00, 17/02, C12Q 1/533, C12N 1/21 //
C07K 103/00, 101/00

(21) International Application Number: PCT/CA02/01453

(22) International Filing Date:
25 September 2002 (25.09.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/324,739 25 September 2001 (25.09.2001) US
60/376,917 1 May 2002 (01.05.2002) US

(71) Applicant (for all designated States except US):
AFFINIUM PHARMACEUTICALS, INC. [CA/CA];
100 University Avenue, 10th Floor, South Tower, Toronto,
Ontario M5J 1V6 (CA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): EDWARDS, Aled
[CA/CA]; 21 Sutherland Drive, Toronto, Ontario M4G
1H1 (CA). DHARAMSI, Akil [CA/CA]; 29 Moresby
Street, Richmond Hill, Ontario L4B 4K9 (CA). VEDADI,

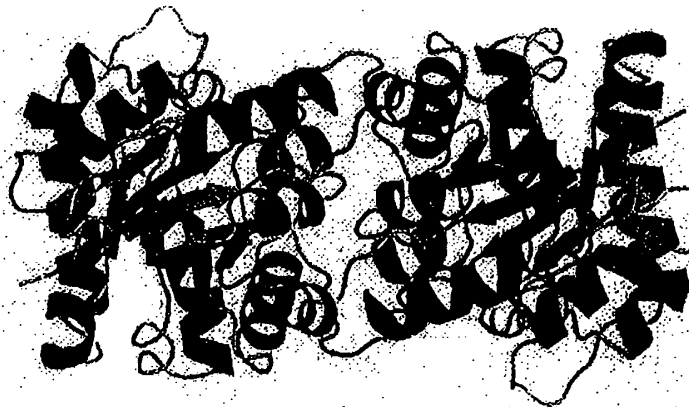
Masoud [CA/CA]; 178 Northwood Drive, Toronto, Ontario M2M 2K3 (CA). ALAM, Muhammad, Zahoor [PK/CA]; #11-440 Malaga Road, Oshawa, Ontario L1J 1N8 (CA). AWREY, Donald [CA/CA]; 2211 Stir Crescent, Mississauga, Ontario L4Y 3V2 (CA). BEATTIE, Bryan [CA/CA]; 2224 Vista Oak Road, Oakville, Ontario L6M 3L7 (CA). HOUSTON, Simon [GB/CA]; 100 Vaughan Road, Apt. 29, Toronto, Ontario M6C 2M1 (CA). KIMBER, Mathew [CA/CA]; 306 Sumach Street, Apt. 7, Toronto, Ontario M5A 3K2 (CA). LAM, Robert [CA/CA]; 85 Thorncliffe Park Drive, Apt. 1505, Toronto, Ontario M4H 1L6 (CA). NETHERY, Kathleen [CA/CA]; 1185 Lansdowne Avenue, Lower Unit, Toronto, Ontario M6H 3Z7 (CA). LI, Qin [CA/CA]; 3455 Rue Durocher, Apt. 314, Montreal, Québec H2X 2C9 (CA). NECAKOV, Sasha [CA/CA]; 530 Brunswick Avenue, Toronto, Ontario M5R 2Z5 (CA). VALLEE, Francois [FR/CA]; 490 Eglinton Avenue East, Apt. 405, Toronto, Ontario M4P 1M4 (CA).

(74) Agent: BERESKIN & PARR; 40 King Street West, 40th Floor, Toronto, Ontario M5H 3Y2 (CA).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,

[Continued on next page]

(54) Title: NOVEL PURIFIED POLYPEPTIDES FROM *PSEUDOMONAS AERUGINOSA*



(57) Abstract: The present invention relates to novel drug targets for pathogenic bacteria. Accordingly, the invention provides purified protein comprising the amino acid sequence set forth in SEQ ID NO: 4. The invention also provides biochemical and biophysical characteristics of the polypeptides of the invention.

WO 03/027274 A2

BEST AVAILABLE COPY



SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VC, VN, YU, ZA, ZM, ZW.

- (84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK,
TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Novel Purified Polypeptides from Pseudomonas aeruginosa

RELATED APPLICATION INFORMATION

This application claims the benefit of priority to the following U.S. Provisional Patent Applications, all of which applications are hereby incorporated by reference in their entirety.

| <i>Provisional Application Number</i> | <i>Filing Date</i> |
|---------------------------------------|--------------------|
| 60/324,739 | September 25, 2001 |
| 60/376,917 | May 1, 2002 |

INTRODUCTION

The discovery of novel antimicrobial agents that work by novel mechanisms is a problem researchers in all fields of drug development face today. The increasing prevalence of drug-resistant pathogens (bacteria, fungi, parasites, etc.) has led to significantly higher mortality rates from infectious diseases and currently presents a serious crisis worldwide. Despite the introduction of second and third generation antimicrobial drugs, certain pathogens have developed resistance to all currently available drugs.

One of the problems contributing to the development of multiple drug resistant pathogens is the limited number of protein targets for antimicrobial drugs. Many of the antibiotics currently in use are structurally related or act through common targets or pathways. Accordingly, adaptive mutation of a single gene may render a pathogenic species resistant to multiple classes of antimicrobial drugs. Therefore, the rapid discovery of drug targets is urgently needed in order to combat the constantly evolving threat by such infectious microorganisms.

Recent advances in bacterial and viral genomics research provides an opportunity for rapid progress in the identification of drug targets. The complete genomic sequences for a number of microorganisms are available. However, knowledge of the complete genomic sequence is only the first step in a long process toward discovery of a viable drug target. The genomic sequence must be annotated to identify open reading frames (ORFs), the essentiality of the protein encoded by the ORF must be determined and the mechanism of action of the gene product must be determined in order to develop a targeted approach to drug discovery.

There are a variety of computer programs available to annotate genomic sequences. Genome annotation involves both identification of genes as well assignment of function thereto based on sequence comparison to homologous proteins with known or predicted functions. However, genome annotation has turned out to be much more of an art than a science. Factors such as splice variants and sequencing errors coupled with the particular algorithms and databases used to annotate the genome can result in significantly different annotations for the same genome. For example, upon reanalysis of the genome of *Mycoplasma pneumoniae* using more rigorous sequence comparisons coupled with molecular biological techniques, such as gel electrophoresis and mass spectrometry, researchers were able to identify several previously unidentified coding sequences, to dismiss a previous identified coding sequence as a likely pseudogene, and to adjust the length of several previously defined ORFs (Dandekar et al. (2000) Nucl. Acids Res. 28(17): 3278-3288). Furthermore, while overall conservation between amino acid sequences generally indicates a conservation of structure and function, specific changes at key residues can lead to significant variation in the biochemical and biophysical properties of a protein. In a comparison of three different functional annotations of the *Mycoplasma genitalium* genome, it was discovered that some genes were assigned three different functions and it was estimated that the overall error rate in the annotations was at least 8% (Brenner (1999) Trends Genet 15(4): 132-3). Accordingly, molecular biological techniques are required to ensure proper genome annotation and identify valid drug targets.

However, confirmation of genome annotation using molecular biological techniques is not an easy proposition due to the unpredictability in expression and purification of polypeptide sequences. Further, in order to carry out structural studies to validate proteins as potential drug targets, it is generally necessary to modify the native proteins in order to facilitate these analyses, e.g., by labeling the protein (e.g., with a heavy atom, isotopic label, polypeptide tag, etc.) or by creating fragments of the polypeptide corresponding to functional domains of a multi-domain protein. Moreover, it is well-known that even small changes in the amino acid sequence of a protein may lead to dramatic affects on protein solubility (Eberstadt et al. (1998) Nature 392: 941-945). Accordingly, genome-wide validation of protein targets will require considerable effort even in light of the sequence of the entire genome of an organism and/or purification conditions for homologs of a particular target.

We have developed reliable, high throughput methods to address some of the shortcomings identified above. In part, using these methods, we have now identified, expressed, and purified a novel antimicrobial target from *Pseudomonas aeruginosa*, or *P. aeruginosa*. Various biophysical, bioinformatic and biochemical studies have been used to characterize the structure and function of the polypeptides of the invention.

TABLE OF CONTENTS

| | |
|---------------------------------------------------------------------|-----|
| RELATED APPLICATION INFORMATION | 1 |
| INTRODUCTION | 1 |
| TABLE OF CONTENTS..... | 3 |
| SUMMARY OF THE INVENTION | 4 |
| BRIEF DESCRIPTION OF THE FIGURES..... | 6 |
| DETAILED DESCRIPTION OF THE INVENTION..... | 9 |
| 1. Definitions..... | 9 |
| 2. Polypeptides of the Invention..... | 26 |
| 3. Nucleic Acids of the Invention | 38 |
| 4. Homology Searching of Nucleotide and Polypeptide Sequences | 47 |
| 5. Analysis of Protein Properties..... | 48 |
| (a) Analysis of Proteins by Mass Spectrometry..... | 48 |
| (b) Analysis of Proteins by Nuclear Magnetic Resonance (NMR) | 50 |
| (c) Analysis of Proteins by X-ray Crystallography..... | 57 |
| (i) X-ray Structure Determination | 57 |
| (ii) X-ray Structure..... | 65 |
| (iii) Structural Equivalents | 66 |
| (iv) Machine Displays and Machine Readable Storage Media | 67 |
| (v) Structurally Similar Molecules and Complexes | 70 |
| (vi) NMR Analysis Using X-ray Structural Data..... | 72 |
| 6. Interacting Proteins | 74 |
| 7. Antibodies | 87 |
| 8. Diagnostic Assays | 90 |
| 9. Drug Discovery..... | 94 |
| (a) Drug Design..... | 95 |
| (b) In Vitro Assays | 104 |
| (c) In Vivo Assays..... | 105 |
| (i) Mouse Soft Tissue Model | 105 |
| (ii) Diffusion Chamber Model..... | 106 |
| (iii) Endocarditis Model..... | 106 |

| | |
|---------------------------------------------------------------------------------------|-----|
| (iv) <i>Osteomyelitis Model</i> | 106 |
| (v) <i>Murine Septic Arthritis Model</i> | 106 |
| (vi) <i>Bacterial Peritonitis Model</i> | 107 |
| 10. <i>Vaccines</i> | 107 |
| 11. <i>Array Analysis</i> | 110 |
| 12. <i>Pharmaceutical Compositions</i> | 113 |
| 13. <i>Antimicrobial Agents</i> | 113 |
| EXEMPLIFICATION..... | 114 |
| EXAMPLE 1 <i>Isolation and Cloning of Nucleic Acid</i> | 115 |
| EXAMPLE 2 <i>Expressing Polypeptides of the Invention</i> | 117 |
| EXAMPLE 3 <i>Expression of Selmet Labeled Polypeptides</i> | 119 |
| EXAMPLE 4 <i>Expression of ¹⁵N Labeled Polypeptides</i> | 120 |
| EXAMPLE 5 <i>Method One for Purifying Polypeptides of the Invention</i> | 121 |
| EXAMPLE 6 <i>Method Two for Purifying Polypeptides of the Invention</i> | 123 |
| EXAMPLE 7 <i>Method Three for Purifying Polypeptides of the Invention</i> | 123 |
| EXAMPLE 8 <i>Mass Spectrometry Analysis via Fingerprint Mapping</i> | 125 |
| EXAMPLE 9 <i>Mass Spectrometry Analysis via High Mass</i> | 127 |
| EXAMPLE 10 <i>Method One for Isolating and Identifying Interacting Proteins</i> | 127 |
| EXAMPLE 11 <i>Method Two for Isolating and Identifying Interacting Proteins</i> | 130 |
| EXAMPLE 12 <i>Sample for Mass Spectrometry of Interacting Proteins</i> | 131 |
| EXAMPLE 13 <i>Mass Spectrometric Analysis of Interacting Proteins</i> | 133 |
| EXAMPLE 14 <i>NMR Analysis</i> | 134 |
| EXAMPLE 15 <i>X-ray Crystallography</i> | 135 |
| EXAMPLE 16 <i>Annotations</i> | 143 |
| EXAMPLE 17 <i>Essential Gene Analysis</i> | 144 |
| EXAMPLE 18 <i>PDB Analysis</i> | 145 |
| EXAMPLE 19 <i>Virtual Genome Analysis</i> | 145 |
| EXAMPLE 20 <i>Epitopic Regions</i> | 146 |
| EQUIVALENTS..... | 146 |
| CLAIMS..... | 150 |

SUMMARY OF THE INVENTION

As part of an effort at genome-wide structural and functional characterization of microbial targets, the present invention provides polypeptides from *P. aeruginosa*. In

various aspects, the invention provides the nucleic acid and amino acid sequences of the polypeptides of the invention. The invention also provides purified, soluble forms of the polypeptides of the invention suitable for structural and functional characterization using a variety of techniques, including, for example, affinity chromatography, mass spectrometry, NMR and x-ray crystallography. The invention further provides modified versions of the polypeptides of the invention to facilitate characterization, including polypeptides labeled with isotopic or heavy atoms and fusion proteins.

A polypeptide of the invention has been crystallized and its structure solved as described in detail below, thereby providing information about the structure of the polypeptide, and druggable regions, domains and the like contained therein, all of which may be used in rational-based drug design efforts.

In general, the biological activity of a polypeptide of the invention is expected to be characterized as having a biochemical activity substantially similar to that of triosephosphate isomerase ("TIM"), having the gene designation of *tpiA*, as described in more detail below. This assignment has been confirmed by solving the X-ray structure of a polypeptide of the invention.

All of the information learned and described herein about the polypeptides of the invention may be used to design modulators of one or more of their biological activities. In particular, information critical to the design of therapeutic and diagnostic molecules, including, for example, the protein domain, druggable regions, structural information, and the like for the polypeptides of the invention is now available or attainable as a result of the ability to prepare, purify and characterize them, and domains, fragments, variants and derivatives thereof.

In other aspects of the invention, structural and functional information about the polypeptides of the invention has and will be obtained. Such information, for example, may be incorporated into databases containing information on the polypeptides of the invention, as well as other polypeptide targets from other microbial species. Such databases will provide investigators with a powerful tool to analyze the polypeptides of the invention and aid in the rapid discovery and design of therapeutic and diagnostic molecules.

In another aspect, modulators, inhibitors, agonists or antagonists against the polypeptides of the invention, or biological complexes containing them, or orthologues thereto, may be used to treat any disease or other treatable condition of a patient (including

humans and animals), and particularly a disease caused by *P. aeruginosa*, such as, for example, one of the following: osteomyelitis, otitis externa, conjunctivitis, keratitis, endophthalmitis, alveolar necrosis, vascular invasion, bacteremia, and burn infection.

The present invention further allows relationships between polypeptides from the same and multiple species to be compared by isolating and studying the various polypeptides of the invention and other proteins. By such comparison studies, which may involve multi-variable analysis as appropriate, it is possible to identify drugs that will affect multiple species or drugs that will affect one or a few species. In such a manner, so-called "wide spectrum" and narrow spectrum" anti-infectives may be identified. Alternatively, drugs that are selective for one or more bacterial or other non-mammalian species, and not for one or more mammalian species (especially human), may be identified (and vice-versa).

In other embodiments, the invention contemplates kits including the subject nucleic acids, polypeptides, crystallized polypeptides, antibodies, and other subject materials, and optionally instructions for their use. Uses for such kits include, for example, diagnostic and therapeutic applications.

The embodiments and practices of the present invention, other embodiments, and their features and characteristics, will be apparent from the description, figures and claims that follow, with all of the claims hereby being incorporated by this reference into this Summary.

BRIEF DESCRIPTION OF THE FIGURES

FIGURE 1 shows the nucleic acid coding sequence for an exemplary polypeptide of the invention as predicted from the genomic sequence of *P. aeruginosa* (SEQ ID NO: 1). This predicted nucleic acid coding sequence was cloned and sequenced to produce the polynucleotide sequence shown in FIGURE 3 (SEQ ID NO: 3).

FIGURE 2 shows the amino acid sequence for an exemplary polypeptide of the invention as predicted from the nucleotide sequence shown in FIGURE 1 (SEQ ID NO: 2).

FIGURE 3 shows the experimentally determined nucleic acid coding sequence for an exemplary polypeptide of the invention (SEQ ID NO: 3).

FIGURE 4 shows the amino acid sequence for the exemplary polypeptide of the invention as predicted from the nucleotide sequence shown in FIGURE 3 (SEQ ID NO: 4).

FIGURE 5 shows the primer sequences used to amplify the nucleic acid of SEQ ID NO: 3. The primers are SEQ ID NO: 5 and SEQ ID NO: 6.

FIGURE 6 contains Table 1, which provides among other things a variety of data and other information on the polypeptides of the invention.

FIGURE 7 contains Table 2, which provides the results of several bioinformatic analyses relating to SEQ ID NO: 2.

FIGURE 8 contains Table 3, which shows information related to the x-ray structure for a polypeptide of the invention as described more fully in EXAMPLE 15.

FIGURE 9 lists the atomic structure coordinates for a polypeptide of the invention derived from x-ray diffraction from a crystal of such polypeptide, as described in more detail in EXAMPLE 15. There are multiple pages to FIGURE 9, labeled 1, 2, 3, etc. The information in such Figure is presented in the following tabular format, with a generic entry provided as an example:

| Record Header | No. | Atom Type | Residue | Residue Number | X | Y | Z | OCC | B |
|---------------|-----|-----------|---------|----------------|-------|--------|--------|-----|-------|
| ATOM 1 | 1 | CB | HIS | 1 | 4.497 | 15.607 | 34.172 | 1 | 70.54 |

In the table, "Record Header" describes the row type, such as "ATOM". "No." refers to the row number. The first "Atom Type" column refers to the atom whose coordinates are measured, with the first letter in the column identifying the atom by its elemental symbol and the subsequent letter defining the location of the atom in the amino acid residue or other molecule. "Residue" and "residue number" identifies the residue of the subject polypeptide. "X, Y, Z" crystallographically define the atomic position of the atom measured. "Occ" is an occupancy factor that refers to the fraction of the molecules in which each atom occupies the position specified by the coordinates. A value of "1" indicates that each atom has the same conformation, i.e., the same position, in all molecules of the crystal. "B" is a thermal factor that is related to the root mean square deviation in the position of the atom around the given atomic coordinate.

FIGURE 10 depicts a clustal V-based sequence alignment of the triosephosphate isomerase protein sequences from six pathogens. The dark shading indicates conserved amino acids across species, with gray areas less conserved.

FIGURE 11 depicts a schematic ribbon diagram of the structure of *P. aeruginosa* triosephosphate isomerase dimer in a side-on view, showing the β barrel architecture. The

shading scheme is used to provide depth perception, with dark representing near objects and light for distant objects.

FIGURE 12 depicts a schematic ribbon diagram of a *P. aeruginosa* triosephosphate isomerase dimer, viewed down the barrel axis.

FIGURE 13 depicts the residues located in and around the catalytic site of yeast TIM complexed with the transition state analogue 2-phosphoglycolic acid (2-PGA). The catalytic residues (E165, H95) at the active site as well as those that anchor the substrate (magenta) are shaded dark and light blue, respectively. The 11-residue flexible loop that appears to "close" the active site upon binding of the substrate to the enzyme is colored red. The residues (A176 and Y208) thought essential for proper closure of the loop are colored green.

FIGURE 14 depicts a schematic ribbon diagram of *P. aeruginosa* TIM illustrating the "open" conformation of the flexible loop, shown here as a dark shaded loop.

FIGURE 15 depicts a schematic ribbon diagram of *S. cerevisiae* TIM complexed with 2-phosphoglycolytic acid (dark spheres), illustrating the "closed" conformation of the flexible loop, shown here as a dark shaded loop.

FIGURE 16 depicts a solvent accessible surface representation of the *P. aeruginosa* TIM dimer, onto which sequence conservation is mapped. The degree of variation in sequence observed among the bacterial organisms *P. aeruginosa*, *E. coli*, *S. aureus*, *S. pneumonia*, *H. pylori* and *E. faecilis* is mapped onto the *P. aeruginosa* TpiA structure. Sequence variation is illustrated by the color coding scheme: completely conserved residues are red with lighter shades of red representing progressively less conservation; white represents residues that have an average degree of sequence conservation; and blue represents residues that are more variable than average. The highly conserved region around the active site (encircled), including the "flexible loop" (arrow), can be clearly identified in this diagram. Additionally, the N-terminal end of the opposing TIM barrel in the dimer is also visible (on the right), showing that this region of the enzyme is less conserved.

FIGURE 17 contains Table 4, which lists the residues in the intersubunit region of TcTIM (*T. cruzi*), PaTIM (*P. aeruginosa*), and HsTIM (*H. sapiens*).

FIGURE 18 depicts potential druggable intersubunit residues in *P. aeruginosa* TIM (PaTIM), Cys43 and Cys88, which are depicted in green.

DETAILED DESCRIPTION OF THE INVENTION

1. Definitions

For convenience, certain terms employed in the specification, examples, and appended claims are collected here. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

The articles "a" and "an" are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

The term "amino acid" is intended to embrace all molecules, whether natural or synthetic, which include both an amino functionality and an acid functionality and capable of being included in a polymer of naturally-occurring amino acids. Exemplary amino acids include naturally-occurring amino acids; analogs, derivatives and congeners thereof; amino acid analogs having variant side chains; and all stereoisomers of any of any of the foregoing.

The term "binding" refers to an association, which may be a stable association, between two molecules, e.g., between a polypeptide of the invention and a binding partner, due to, for example, electrostatic, hydrophobic, ionic and/or hydrogen-bond interactions under physiological conditions.

A "comparison window," as used herein, refers to a conceptual segment of at least 20 contiguous amino acid positions wherein a protein sequence may be compared to a reference sequence of at least 20 contiguous amino acids and wherein the portion of the protein sequence in the comparison window may comprise additions or deletions (i.e., gaps) of 20 percent or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2: 482, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48: 443, by the search

for similarity method of Pearson and Lipman (1988) Proc. Natl. Acad. Sci. (U.S.A.) 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection, and the best alignment (i.e., resulting in the highest percentage of homology over the comparison window) generated by the various methods may be identified.

The term "complex" refers to an association between at least two moieties (e.g. chemical or biochemical) that have an affinity for one another. Examples of complexes include associations between antigen/antibodies, lectin/avidin, target polynucleotide/probe oligonucleotide, antibody/anti-antibody, receptor/ligand, enzyme/ligand, polypeptide/polypeptide, polypeptide/polynucleotide, polypeptide/co-factor, polypeptide/substrate, polypeptide/inhibitor, polypeptide/small molecule, and the like. "Member of a complex" refers to one moiety of the complex, such as an antigen or ligand. "Protein complex" or "polypeptide complex" refers to a complex comprising at least one polypeptide.

The term "conserved residue" refers to an amino acid that is a member of a group of amino acids having certain common properties. The term "conservative amino acid substitution" refers to the substitution (conceptually or otherwise) of an amino acid from one such group with a different amino acid from the same group. A functional way to define common properties between individual amino acids is to analyze the normalized frequencies of amino acid changes between corresponding proteins of homologous organisms (Schulz, G. E. and R. H. Schirmer., Principles of Protein Structure, Springer-Verlag). According to such analyses, groups of amino acids may be defined where amino acids within a group exchange preferentially with each other, and therefore resemble each other most in their impact on the overall protein structure (Schulz, G. E. and R. H. Schirmer, Principles of Protein Structure, Springer-Verlag). One example of a set of amino acid groups defined in this manner include: (i) a charged group, consisting of Glu and Asp, Lys, Arg and His, (ii) a positively-charged group, consisting of Lys, Arg and His, (iii) a negatively-charged group, consisting of Glu and Asp, (iv) an aromatic group, consisting of Phe, Tyr and Trp, (v) a nitrogen ring group, consisting of His and Trp, (vi) a large aliphatic nonpolar group, consisting of Val, Leu and Ile, (vii) a slightly-polar group, consisting of Met and Cys, (viii) a small-residue group, consisting of Ser, Thr, Asp, Asn, Gly, Ala, Glu, Gln and Pro, (ix) an aliphatic group consisting of Val, Leu, Ile, Met and Cys, and (x) a small hydroxyl group consisting of Ser and Thr.

The term "domain", when used in connection with a polypeptide, refers to a specific region within such polypeptide that comprises a particular structure or mediates a particular function. In the typical case, a domain of a polypeptide of the invention is a fragment of the polypeptide. In certain instances, a domain is a structurally stable domain, as evidenced, for example, by mass spectroscopy, or by the fact that a modulator may bind to a druggable region of the domain.

The term "druggable region", when used in reference to a polypeptide, nucleic acid, complex and the like, refers to a region of the molecule which is a target or is a likely target for binding a modulator. For a polypeptide, a druggable region generally refers to a region wherein several amino acids of a polypeptide would be capable of interacting with a modulator or other molecule. For a polypeptide or complex thereof, exemplary druggable regions including binding pockets and sites, enzymatic active sites, interfaces between domains of a polypeptide or complex, surface grooves or contours or surfaces of a polypeptide or complex which are capable of participating in interactions with another molecule. In certain instances, the interacting molecule is another polypeptide, which may be naturally-occurring. In other instances, the druggable region is on the surface of the molecule.

Druggable regions may be described and characterized in a number of ways. For example, a druggable region may be characterized by some or all of the amino acids that make up the region, or the backbone atoms thereof, or the side chain atoms thereof (optionally with or without the C α atoms). Alternatively, in certain instances, the volume of a druggable region corresponds to that of a carbon based molecule of at least about 200 amu and often up to about 800 amu. In other instances, it will be appreciated that the volume of such region may correspond to a molecule of at least about 600 amu and often up to about 1600 amu or more.

Alternatively, a druggable region may be characterized by comparison to other regions on the same or other molecules. For example, the term "affinity region" refers to a druggable region on a molecule (such as a polypeptide of the invention) that is present in several other molecules, in so much as the structures of the same affinity regions are sufficiently the same so that they are expected to bind the same or related structural analogs. An example of an affinity region is an ATP-binding site of a protein kinase that is found in several protein kinases (whether or not of the same origin). The term "selectivity

region" refers to a druggable region of a molecule that may not be found on other molecules, in so much as the structures of different selectivity regions are sufficiently different so that they are not expected to bind the same or related structural analogs. An exemplary selectivity region is a catalytic domain of a protein kinase that exhibits specificity for one substrate. In certain instances, a single modulator may bind to the same affinity region across a number of proteins that have a substantially similar biological function, whereas the same modulator may bind to only one selectivity region of one of those proteins.

Continuing with examples of different druggable regions, the term "undesired region" refers to a druggable region of a molecule that upon interacting with another molecule results in an undesirable affect. For example, a binding site that oxidizes the interacting molecule (such as P-450 activity) and thereby results in increased toxicity for the oxidized molecule may be deemed a "undesired region". Other examples of potential undesired regions includes regions that upon interaction with a drug decrease the membrane permeability of the drug, increase the excretion of the drug, or increase the blood brain transport of the drug. It may be the case that, in certain circumstances, an undesired region will be no longer be deemed an undesired region because the affect of the region will be favorable, e.g., a drug intended to treat a brain condition would benefit from interacting with a region that resulted in increased blood brain transport, whereas the same region could be deemed undesirable for drugs that were not intended to be delivered to the brain.

When used in reference to a druggable region, the "selectivity" or "specificity" of a molecule such as a modulator to a druggable region may be used to describe the binding between the molecule and a druggable region. For example, the selectivity of a modulator with respect to a druggable region may be expressed by comparison to another modulator, using the respective values of K_d (i.e., the dissociation constants for each modulator-druggable region complex) or, in cases where a biological effect is observed below the K_d , the ratio of the respective EC_{50} 's (i.e., the concentrations that produce 50% of the maximum response for the modulator interacting with each druggable region).

A "fusion protein" or "fusion polypeptide" refers to a chimeric protein as that term is known in the art and may be constructed using methods known in the art. In many examples of fusion proteins, there are two different polypeptide sequences, and in certain cases, there may be more. The sequences may be linked in frame. A fusion protein may include a domain which is found (albeit in a different protein) in an organism which also

expresses the first protein, or it may be an "interspecies", "intergenic", etc. fusion expressed by different kinds of organisms. In various embodiments, the fusion polypeptide may comprise one or more amino acid sequences linked to a first polypeptide. In the case where more than one amino acid sequence is fused to a first polypeptide, the fusion sequences may be multiple copies of the same sequence, or alternatively, may be different amino acid sequences. The fusion polypeptides may be fused to the N-terminus, the C-terminus, or the N- and C-terminus of the first polypeptide. Exemplary fusion proteins include polypeptides comprising a glutathione S-transferase tag (GST-tag), histidine tag (His-tag), an immunoglobulin domain or an immunoglobulin binding domain.

The term "gene" refers to a nucleic acid comprising an open reading frame encoding a polypeptide having exon sequences and optionally intron sequences. The term "intron" refers to a DNA sequence present in a given gene which is not translated into protein and is generally found between exons.

The term "having substantially similar biological activity", when used in reference to two polypeptides, refers to a biological activity of a first polypeptide which is substantially similar to at least one of the biological activities of a second polypeptide. A substantially similar biological activity means that the polypeptides carry out a similar function, e.g., a similar enzymatic reaction or a similar physiological process, etc. For example, two homologous proteins may have a substantially similar biological activity if they are involved in a similar enzymatic reaction, e.g., they are both kinases which catalyze phosphorylation of a substrate polypeptide, however, they may phosphorylate different regions on the same protein substrate or different substrate proteins altogether. Alternatively, two homologous proteins may also have a substantially similar biological activity if they are both involved in a similar physiological process, e.g., transcription. For example, two proteins may be transcription factors, however, they may bind to different DNA sequences or bind to different polypeptide interactors. Substantially similar biological activities may also be associated with proteins carrying out a similar structural role, for example, two membrane proteins.

The term "isolated polypeptide" refers to a polypeptide, in certain embodiments prepared from recombinant DNA or RNA, or of synthetic origin, or some combination thereof, which (1) is not associated with proteins that it is normally found with in nature, (2) is isolated from the cell in which it normally occurs, (3) is isolated free of other proteins

from the same cellular source, e.g. free of *P. aeruginosa* proteins, (4) is expressed by a cell from a different species, or (5) does not occur in nature.

The term "isolated nucleic acid" refers to a polynucleotide of genomic, cDNA, or synthetic origin or some combination thereof, which (1) is not associated with the cell in which the "isolated nucleic acid" is found in nature, or (2) is operably linked to a polynucleotide to which it is not linked in nature.

The terms "label" or "labeled" refer to incorporation or attachment, optionally covalently or non-covalently, of a detectable marker into a molecule, such as a polypeptide. Various methods of labeling polypeptides are known in the art and may be used. Examples of labels for polypeptides include, but are not limited to, the following: radioisotopes, fluorescent labels, heavy atoms, enzymatic labels or reporter genes, chemiluminescent groups, biotinyl groups, predetermined polypeptide epitopes recognized by a secondary reporter (e.g., leucine zipper pair sequences, binding sites for secondary antibodies, metal binding domains, epitope tags). Examples and use of such labels are described in more detail below. In some embodiments, labels are attached by spacer arms of various lengths to reduce potential steric hindrance.

The term "mammal" is known in the art, and exemplary mammals include humans, primates, bovines, porcines, canines, felines, and rodents (e.g., mice and rats).

The term "modulation", when used in reference to a functional property or biological activity or process (e.g., enzyme activity or receptor binding), refers to the capacity to either up regulate (e.g., activate or stimulate), down regulate (e.g., inhibit or suppress) or otherwise change a quality of such property, activity or process. In certain instances, such regulation may be contingent on the occurrence of a specific event, such as activation of a signal transduction pathway, and/or may be manifest only in particular cell types.

The term "modulator" refers to a polypeptide, nucleic acid, macromolecule, complex, molecule, small molecule, compound, species or the like (naturally-occurring or non-naturally-occurring), or an extract made from biological materials such as bacteria, plants, fungi, or animal cells or tissues, that may be capable of causing modulation. Modulators may be evaluated for potential activity as inhibitors or activators (directly or indirectly) of a functional property, biological activity or process, or combination of them, (e.g., agonist, partial antagonist, partial agonist, inverse agonist, antagonist, anti-microbial

agents, inhibitors of microbial infection or proliferation, and the like) by inclusion in assays. In such assays, many modulators may be screened at one time. The activity of a modulator may be known, unknown or partially known.

The term "motif" refers to an amino acid sequence that is commonly found in a protein of a particular structure or function. Typically, a consensus sequence is defined to represent a particular motif. The consensus sequence need not be strictly defined and may contain positions of variability, degeneracy, variability of length, etc. The consensus sequence may be used to search a database to identify other proteins that may have a similar structure or function due to the presence of the motif in its amino acid sequence. For example, on-line databases may be searched with a consensus sequence in order to identify other proteins containing a particular motif. Various search algorithms and/or programs may be used, including FASTA, BLAST or ENTREZ. FASTA and BLAST are available as a part of the GCG sequence analysis package (University of Wisconsin, Madison, Wis.). ENTREZ is available through the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD.

The term "naturally-occurring", as applied to an object, refers to the fact that an object may be found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism (including bacteria) that may be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is naturally-occurring.

The term "nucleic acid" refers to a polymeric form of nucleotides, either ribonucleotides or deoxynucleotides or a modified form of either type of nucleotide. The terms should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single-stranded (such as sense or antisense) and double-stranded polynucleotides.

The term "nucleic acid of the invention" refers to a nucleic acid encoding a polypeptide of the invention, e.g., a nucleic acid comprising a sequence consisting of, or consisting essentially of, the polynucleotide sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 3. A nucleic acid of the invention may comprise all, or a portion of: the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3; a nucleotide sequence at least 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 1 or SEQ ID NO: 3; a nucleotide sequence that hybridizes under stringent conditions to SEQ ID NO: 1 or SEQ ID

NO: 3; nucleotide sequences encoding polypeptides that are functionally equivalent to polypeptides of the invention; nucleotide sequences encoding polypeptides at least about 60%, 70%, 80%, 85%, 90%, 95%, 98%, 99% homologous or identical with an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4; nucleotide sequences encoding polypeptides having an activity of a polypeptide of the invention and having at least about 60%, 70%, 80%, 85%, 90%, 95%, 98%, 99% or more homology or identity with SEQ ID NO: 2 or SEQ ID NO: 4; nucleotide sequences that differ by 1 to about 2, 3, 5, 7, 10, 15, 20, 30, 50, 75 or more nucleotide substitutions, additions or deletions, such as allelic variants, of SEQ ID NO: 1 and SEQ ID NO: 3; nucleic acids derived from and evolutionarily related to SEQ ID NO: 1 or SEQ ID NO: 3; and complements of, and nucleotide sequences resulting from the degeneracy of the genetic code, for all of the foregoing and other nucleic acids of the invention. Nucleic acids of the invention also include homologs, e.g., orthologs and paralogs, of SEQ ID NO: 1 or SEQ ID NO: 3 and also variants of SEQ ID NO: 1 or SEQ ID NO: 3 which have been codon optimized for expression in a particular organism (e.g., host cell).

The term "operably linked", when describing the relationship between two nucleic acid regions, refers to a juxtaposition wherein the regions are in a relationship permitting them to function in their intended manner. For example, a control sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences, such as when the appropriate molecules (e.g., inducers and polymerases) are bound to the control or regulatory sequence(s).

The term "phenotype" refers to the entire physical, biochemical, and physiological makeup of a cell, e.g., having any one trait or any group of traits.

The term "polypeptide", and the terms "protein" and "peptide" which are used interchangeably herein, refers to a polymer of amino acids. Exemplary polypeptides include gene products, naturally-occurring proteins, homologs, orthologs, paralogs, fragments, and other equivalents, variants and analogs of the foregoing.

The terms "polypeptide fragment" or "fragment", when used in reference to a reference polypeptide, refers to a polypeptide in which amino acid residues are deleted as compared to the reference polypeptide itself, but where the remaining amino acid sequence is usually identical to the corresponding positions in the reference polypeptide. Such

deletions may occur at the amino-terminus or carboxy-terminus of the reference polypeptide, or alternatively both. Fragments typically are at least 5, 6, 8 or 10 amino acids long, at least 14 amino acids long, at least 20, 30, 40 or 50 amino acids long, at least 75 amino acids long, or at least 100, 150, 200, 300, 500 or more amino acids long. A fragment can retain one or more of the biological activities of the reference polypeptide. In certain embodiments, a fragment may comprise a druggable region, and optionally additional amino acids on one or both sides of the druggable region, which additional amino acids may number from 5, 10, 15, 20, 30, 40, 50, or up to 100 or more residues. Further, fragments can include a sub-fragment of a specific region, which sub-fragment retains the function of the region from which it is derived. In another embodiment, a fragment may have immunogenic properties.

The term "polypeptide of the invention" refers to a polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4, or an equivalent or fragment thereof, e.g., a polypeptide comprising a sequence consisting of, or consisting essentially of, the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4. Polypeptides of the invention include polypeptides comprising all or a portion of the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4 with 1 to about 2, 3, 5, 7, 10, 15, 20, 30, 50, 75 or more conservative amino acid substitutions; an amino acid sequence that is at least 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 2 or SEQ ID NO: 4; and functional fragments thereof. Polypeptides of the invention also include homologs, e.g., orthologs and paralogs, of SEQ ID NO: 2 or SEQ ID NO: 4.

The term "purified" refers to an object species that is the predominant species present (i.e., on a molar basis it is more abundant than any other individual species in the composition). A "purified fraction" is a composition wherein the object species comprises at least about 50 percent (on a molar basis) of all species present. In making the determination of the purity of a species in solution or dispersion, the solvent or matrix in which the species is dissolved or dispersed is usually not included in such determination; instead, only the species (including the one of interest) dissolved or dispersed are taken into account. Generally, a purified composition will have one species that comprises more than about 80 percent of all species present in the composition, more than about 85%, 90%, 95%, 99% or more of all species present. The object species may be purified to essential homogeneity (contaminant species cannot be detected in the composition by conventional

detection methods) wherein the composition consists essentially of a single species. A skilled artisan may purify a polypeptide of the invention using standard techniques for protein purification in light of the teachings herein. Purity of a polypeptide may be determined by a number of methods known to those of skill in the art, including for example, amino-terminal amino acid sequence analysis, gel electrophoresis, mass-spectrometry analysis and the methods described in the Exemplification section herein.

The terms "recombinant protein" or "recombinant polypeptide" refer to a polypeptide which is produced by recombinant DNA techniques. An example of such techniques includes the case when DNA encoding the expressed protein is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the protein or polypeptide encoded by the DNA.

A "reference sequence" is a defined sequence used as a basis for a sequence comparison; a reference sequence may be a subset of a larger sequence, for example, as a segment of a full-length protein given in a sequence listing such as SEQ ID NO: 2 or SEQ ID NO: 4, or may comprise a complete protein sequence. Generally, a reference sequence is at least 200, 300 or 400 nucleotides in length, frequently at least 600 nucleotides in length, and often at least 800 nucleotides in length (or the protein equivalent if it is shorter or longer in length). Because two proteins may each (1) comprise a sequence (i.e., a portion of the complete protein sequence) that is similar between the two proteins, and (2) may further comprise a sequence that is divergent between the two proteins, sequence comparisons between two (or more) proteins are typically performed by comparing sequences of the two proteins over a "comparison window" to identify and compare local regions of sequence similarity.

The term "regulatory sequence" is a generic term used throughout the specification to refer to polynucleotide sequences, such as initiation signals, enhancers, regulators and promoters, that are necessary or desirable to affect the expression of coding and non-coding sequences to which they are operably linked. Exemplary regulatory sequences are described in Goeddel; *Gene Expression Technology. Methods in Enzymology*, Academic Press, San Diego, CA (1990), and include, for example, the early and late promoters of SV40, adenovirus or cytomegalovirus immediate early promoter, the lac system, the trp system, the TAC or TRC system, T7 promoter whose expression is directed by T7 RNA polymerase, the major operator and promoter regions of phage lambda, the control regions for fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes,

the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α -mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof. The nature and use of such control sequences may differ depending upon the host organism. In prokaryotes, such regulatory sequences generally include promoter, ribosomal binding site, and transcription termination sequences. The term "regulatory sequence" is intended to include, at a minimum, components whose presence may influence expression, and may also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences. In certain embodiments, transcription of a polynucleotide sequence is under the control of a promoter sequence (or other regulatory sequence) which controls the expression of the polynucleotide in a cell-type in which expression is intended. It will also be understood that the polynucleotide can be under the control of regulatory sequences which are the same or different from those sequences which control expression of the naturally-occurring form of the polynucleotide.

The term "reporter gene" refers to a nucleic acid comprising a nucleotide sequence encoding a protein that is readily detectable either by its presence or activity, including, but not limited to, luciferase, fluorescent protein (e.g., green fluorescent protein), chloramphenicol acetyl transferase, β -galactosidase, secreted placental alkaline phosphatase, β -lactamase, human growth hormone, and other secreted enzyme reporters. Generally, a reporter gene encodes a polypeptide not otherwise produced by the host cell, which is detectable by analysis of the cell(s), e.g., by the direct fluorometric, radioisotopic or spectrophotometric analysis of the cell(s) and preferably without the need to kill the cells for signal analysis. In certain instances, a reporter gene encodes an enzyme, which produces a change in fluorometric properties of the host cell, which is detectable by qualitative, quantitative or semiquantitative function or transcriptional activation. Exemplary enzymes include esterases, β -lactamase, phosphatases, peroxidases, proteases (tissue plasminogen activator or urokinase) and other enzymes whose function may be detected by appropriate chromogenic or fluorogenic substrates known to those skilled in the art or developed in the future.

The term "sequence homology" refers to the proportion of base matches between two nucleic acid sequences or the proportion of amino acid matches between two amino acid sequences. When sequence homology is expressed as a percentage, e.g., 50%, the

percentage denotes the proportion of matches over the length of sequence from a desired sequence (e.g., SEQ. ID NO: 1) that is compared to some other sequence. Gaps (in either of the two sequences) are permitted to maximize matching; gap lengths of 15 bases or less are usually used, 6 bases or less are used more frequently, with 2 bases or less used even more frequently. The term "sequence identity" means that sequences are identical (i.e., on a nucleotide-by-nucleotide basis for nucleic acids or amino acid-by-amino acid basis for polypeptides) over a window of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over the comparison window, determining the number of positions at which the identical amino acids occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the comparison window, and multiplying the result by 100 to yield the percentage of sequence identity. Methods to calculate sequence identity are known to those of skill in the art and described in further detail below.

The term "small molecule" refers to a compound, which has a molecular weight of less than about 5 kD, less than about 2.5 kD, less than about 1.5 kD, or less than about 0.9 kD. Small molecules may be, for example, nucleic acids, peptides, polypeptides, peptide nucleic acids, peptidomimetics, carbohydrates, lipids or other organic (carbon containing) or inorganic molecules. Many pharmaceutical companies have extensive libraries of chemical and/or biological mixtures, often fungal, bacterial, or algal extracts, which can be screened with any of the assays of the invention. The term "small organic molecule" refers to a small molecule that is often identified as being an organic or medicinal compound, and does not include molecules that are exclusively nucleic acids, peptides or polypeptides.

The term "soluble" as used herein with reference to a polypeptide of the invention or other protein, means that upon expression in cell culture, at least some portion of the polypeptide or protein expressed remains in the cytoplasmic fraction of the cell and does not fractionate with the cellular debris upon lysis and centrifugation of the lysate. Solubility of a polypeptide may be increased by a variety of art recognized methods, including fusion to a heterologous amino acid sequence, deletion of amino acid residues, amino acid substitution (e.g., enriching the sequence with amino acid residues having hydrophilic side chains), and chemical modification (e.g., addition of hydrophilic groups). The solubility of polypeptides may be measured using a variety of art recognized techniques, including, dynamic light scattering to determine aggregation state, UV absorption, centrifugation to separate aggregated from non-aggregated material, and SDS

gel electrophoresis (e.g., the amount of protein in the soluble fraction is compared to the amount of protein in the soluble and insoluble fractions combined). When expressed in a host cell, the polypeptides of the invention may be at least about 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more soluble, e.g., at least about 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more of the total amount of protein expressed in the cell is found in the cytoplasmic fraction. In certain embodiments, a one liter culture of cells expressing a polypeptide of the invention will produce at least about 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 30, 40, 50 milligrams or more of soluble protein. In an exemplary embodiment, a polypeptide of the invention is at least about 10% soluble and will produce at least about 1 milligram of protein from a one liter cell culture.

The term "specifically hybridizes" refers to detectable and specific nucleic acid binding. Polynucleotides, oligonucleotides and nucleic acids of the invention selectively hybridize to nucleic acid strands under hybridization and wash conditions that minimize appreciable amounts of detectable binding to nonspecific nucleic acids. Stringent conditions may be used to achieve selective hybridization conditions as known in the art and discussed herein. Generally, the nucleic acid sequence homology between the polynucleotides, oligonucleotides, and nucleic acids of the invention and a nucleic acid sequence of interest will be at least 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 98%, 99%, or more. In certain instances, hybridization and washing conditions are performed at under stringent conditions according to conventional hybridization procedures and as described further herein.

The terms "stringent conditions" or "stringent hybridization conditions" refer to conditions which promote specific hybridization between two complementary polynucleotide strands so as to form a duplex. Stringent conditions may be selected to be about 5°C lower than the thermal melting point (T_m) for a given polynucleotide duplex at a defined ionic strength and pH. The length of the complementary polynucleotide strands and their GC content will determine the T_m of the duplex, and thus the hybridization conditions necessary for obtaining a desired specificity of hybridization. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the a polynucleotide sequence hybridizes to a perfectly matched complementary strand. In certain cases it may be desirable to increase the stringency of the hybridization conditions to be about equal to the T_m for a particular duplex.

A variety of techniques for estimating the T_m are available. Typically, G-C base pairs in a duplex are estimated to contribute about 3°C to the T_m , while A-T base pairs are estimated to contribute about 2°C, up to a theoretical maximum of about 80-100°C. However, more sophisticated models of T_m are available in which G-C stacking interactions, solvent effects, the desired assay temperature and the like are taken into account. For example, probes can be designed to have a dissociation temperature (T_d) of approximately 60°C, using the formula: $T_d = (((((3 \times \#GC) + (2 \times \#AT)) \times 37) - 562) / \#bp) - 5$; where $\#GC$, $\#AT$, and $\#bp$ are the number of guanine-cytosine base pairs, the number of adenine-thymine base pairs, and the number of total base pairs, respectively, involved in the formation of the duplex.

Hybridization may be carried out in 5xSSC, 4xSSC, 3xSSC, 2xSSC, 1xSSC or 0.2xSSC for at least about 1 hour, 2 hours, 5 hours, 12 hours, or 24 hours. The temperature of the hybridization may be increased to adjust the stringency of the reaction, for example, from about 25°C (room temperature), to about 45°C, 50°C, 55°C, 60°C, or 65°C. The hybridization reaction may also include another agent affecting the stringency, for example, hybridization conducted in the presence of 50% formamide increases the stringency of hybridization at a defined temperature.

The hybridization reaction may be followed by a single wash step, or two or more wash steps, which may be at the same or a different salinity and temperature. For example, the temperature of the wash may be increased to adjust the stringency from about 25°C (room temperature), to about 45°C, 50°C, 55°C, 60°C, 65°C, or higher. The wash step may be conducted in the presence of a detergent, e.g., 0.1 or 0.2% SDS. For example, hybridization may be followed by two wash steps at 65°C each for about 20 minutes in 2xSSC, 0.1% SDS, and optionally two additional wash steps at 65°C each for about 20 minutes in 0.2xSSC, 0.1%SDS.

Exemplary stringent hybridization conditions include overnight hybridization at 65°C in a solution comprising, or consisting of, 50% formamide, 10xDenhardt (0.2% Ficoll, 0.2% Polyvinylpyrrolidone, 0.2% bovine serum albumin) and 200 µg/ml of denatured carrier DNA, e.g., sheared salmon sperm DNA, followed by two wash steps at 65°C each for about 20 minutes in 2xSSC, 0.1% SDS, and two wash steps at 65°C each for about 20 minutes in 0.2xSSC, 0.1%SDS.

Hybridization may consist of hybridizing two nucleic acids in solution, or a nucleic acid in solution to a nucleic acid attached to a solid support, e.g., a filter. When one nucleic

acid is on a solid support, a prehybridization step may be conducted prior to hybridization. Prehybridization may be carried out for at least about 1 hour, 3 hours or 10 hours in the same solution and at the same temperature as the hybridization solution (without the complementary polynucleotide strand).

Appropriate stringency conditions are known to those skilled in the art or may be determined experimentally by the skilled artisan. See, for example, *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-12.3.6; Sambrook et al., 1989, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, N.Y.; S. Agrawal (ed.) *Methods in Molecular Biology*, volume 20; Tijssen (1993) *Laboratory Techniques in biochemistry and molecular biology-hybridization with nucleic acid probes*, e.g., part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York; and Tibanyenda, N. et al., *Eur. J. Biochem.* 139:19 (1984) and Ebel, S. et al., *Biochem.* 31:12083 (1992).

As applied to proteins, the term "substantial identity" means that two protein sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap weights, typically share at least about 70 percent sequence identity, alternatively at least about 80, 85, 90, 95 percent sequence identity or more. In certain instances, residue positions that are not identical differ by conservative amino acid substitutions, which are described above.

The term "structural motif", when used in reference to a polypeptide, refers to a polypeptide that, although it may have different amino acid sequences, may result in a similar structure, wherein by structure is meant that the motif forms generally the same tertiary structure, or that certain amino acid residues within the motif, or alternatively their backbone or side chains (which may or may not include the C α atoms of the side chains) are positioned in a like relationship with respect to one another in the motif.

The term "test compound" refers to a molecule to be tested by one or more screening method(s) as a putative modulator of a polypeptide of the invention or other biological entity or process. A test compound is usually not known to bind to a target of interest. The term "control test compound" refers to a compound known to bind to the target (e.g., a known agonist, antagonist, partial agonist or inverse agonist). The term "test compound" does not include a chemical added as a control condition that alters the function of the target to determine signal specificity in an assay. Such control chemicals or conditions include chemicals that 1) nonspecifically or substantially disrupt protein

structure (e.g., denaturing agents (e.g., urea or guanidinium), chaotropic agents, sulfhydryl reagents (e.g., dithiothreitol and β -mercaptoethanol), and proteases), 2) generally inhibit cell metabolism (e.g., mitochondrial uncouplers) and 3) non-specifically disrupt electrostatic or hydrophobic interactions of a protein (e.g., high salt concentrations, or detergents at concentrations sufficient to non-specifically disrupt hydrophobic interactions). Further, the term "test compound" also does not include compounds known to be unsuitable for a therapeutic use for a particular indication due to toxicity of the subject. In certain embodiments, various predetermined concentrations of test compounds are used for screening such as 0.01 μ M, 0.1 μ M, 1.0 μ M, and 10.0 μ M. Examples of test compounds include, but are not limited to, peptides, nucleic acids, carbohydrates, and small molecules. The term "novel test compound" refers to a test compound that is not in existence as of the filing date of this application. In certain assays using novel test compounds, the novel test compounds comprise at least about 50%, 75%, 85%, 90%, 95% or more of the test compounds used in the assay or in any particular trial of the assay.

The term "therapeutically effective amount" refers to that amount of a modulator, drug or other molecule which is sufficient to effect treatment when administered to a subject in need of such treatment. The therapeutically effective amount will vary depending upon the subject and disease condition being treated, the weight and age of the subject, the severity of the disease condition, the manner of administration and the like, which can readily be determined by one of ordinary skill in the art.

The term "transfection" means the introduction of a nucleic acid, e.g., an expression vector, into a recipient cell, which in certain instances involves nucleic acid-mediated gene transfer. The term "transformation" refers to a process in which a cell's genotype is changed as a result of the cellular uptake of exogenous nucleic acid. For example, a transformed cell may express a recombinant form of a polypeptide of the invention or antisense expression may occur from the transferred gene so that the expression of a naturally-occurring form of the gene is disrupted.

The term "transgene" means a nucleic acid sequence, which is partly or entirely heterologous to a transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs

from that of the natural gene or its insertion results in a knockout). A transgene may include one or more regulatory sequences and any other nucleic acids, such as introns, that may be necessary for optimal expression.

The term "transgenic animal" refers to any animal, for example, a mouse, rat or other non-human mammal, a bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA. In the typical transgenic animals described herein, the transgene causes cells to express a recombinant form of a protein. However, transgenic animals in which the recombinant gene is silent are also contemplated.

The term "vector" refers to a nucleic acid capable of transporting another nucleic acid to which it has been linked. One type of vector which may be used in accord with the invention is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Other vectors include those capable of autonomous replication and expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer to circular double stranded DNA molecules which, in their vector form are not bound to the chromosome. In the present specification, "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

Unless otherwise indicated, all numbers expressing quantities of ingredients, reaction conditions, and so forth used in the specification and claims are to be understood as being modified in all instances by the term "about." Accordingly, unless indicated to the contrary, the numerical parameters set forth in this specification and attached claims are approximations that may vary depending upon the desired properties sought to be obtained by the present invention.

2. Polypeptides of the Invention

The present invention makes available in a variety of embodiments soluble, purified and/or isolated forms of the polypeptides of the invention. Milligram quantities of an exemplary polypeptide of the invention, SEQ ID NO: 4 (optionally with a tag, and optionally labeled), have been isolated in a highly purified form. The present invention provides for expressing and purifying polypeptides of the invention in quantities that equal or exceed the quantity of polypeptide(s) of the invention expressed and purified as provided in the Exemplification section below (or smaller amount(s) thereof, such as 25%, 33%, 50% or 75% of the amount(s) so expressed and/or purified).

In one aspect, the present invention contemplates an isolated polypeptide comprising (a) the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4, (b) the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4 with 1 to about 20 conservative amino acid substitutions, deletions or additions, (c) an amino acid sequence that is at least 90% identical to SEQ ID NO: 2 or SEQ ID NO: 4 or (d) a functional fragment of a polypeptide having an amino acid sequence set forth in (a), (b) or (c). In another aspect, the present invention contemplates a composition comprising such an isolated polypeptide and less than about 10%, or alternatively 5%, or alternatively 1%, contaminating biological macromolecules or polypeptides.

It may be the case that the amino acid sequence of SEQ ID NO: 4 differs from that of SEQ ID NO: 2 by one or more amino acids. SEQ ID NO: 4 is determined from the experimentally determined nucleic acid sequence SEQ ID NO: 3, and SEQ ID NO: 2 is determined from SEQ ID NO: 1, which is obtained as described in EXAMPLE 1. In such a case, the present invention contemplates the specific amino acid sequences of SEQ ID NO: 2 and SEQ ID NO: 4, and variants thereof, as well as any differences (if any) in the polypeptides of the invention based on those SEQ ID NOS and nucleic acid sequences encoding the same.

In certain embodiments, a polypeptide of the invention is a fusion protein containing a domain which increases its solubility and/or facilitates its purification, identification, detection, and/or structural characterization. Exemplary domains, include, for example, glutathione S-transferase (GST), protein A, protein G, calmodulin-binding peptide, thioredoxin, maltose binding protein, HA, myc, poly arginine, poly His, poly His-Asp or

FLAG fusion proteins and tags. Additional exemplary domains include domains that alter protein localization *in vivo*, such as signal peptides, type III secretion system-targeting peptides, transcytosis domains, nuclear localization signals, etc. In various embodiments, a polypeptide of the invention may comprise one or more heterologous fusions. Polypeptides may contain multiple copies of the same fusion domain or may contain fusions to two or more different domains. The fusions may occur at the N-terminus of the polypeptide, at the C-terminus of the polypeptide, or at both the N- and C-terminus of the polypeptide. It is also within the scope of the invention to include linker sequences between a polypeptide of the invention and the fusion domain in order to facilitate construction of the fusion protein or to optimize protein expression or structural constraints of the fusion protein. In another embodiment, the polypeptide may be constructed so as to contain protease cleavage sites between the fusion polypeptide and polypeptide of the invention in order to remove the tag after protein expression or thereafter. Examples of suitable endoproteases, include, for example, Factor Xa and TEV proteases.

In another embodiment, a polypeptide of the invention may be modified so that its rate of traversing the cellular membrane is increased. For example, the polypeptide may be fused to a second peptide which promotes "transcytosis," e.g., uptake of the peptide by cells. The peptide may be a portion of the HIV transactivator (TAT) protein, such as the fragment corresponding to residues 37-62 or 48-60 of TAT, portions which have been observed to be rapidly taken up by a cell *in vitro* (Green and Loewenstein, (1989) Cell 55:1179-1188). Alternatively, the internalizing peptide may be derived from the *Drosophila antennapedia* protein, or homologs thereof. The 60 amino acid long homeodomain of the homeo-protein antennapedia has been demonstrated to translocate through biological membranes and can facilitate the translocation of heterologous polypeptides to which it is coupled. Thus, polypeptides may be fused to a peptide consisting of about amino acids 42-58 of *Drosophila antennapedia* or shorter fragments for transcytosis (Derossi et al. (1996) J Biol Chem 271:18188-18193; Derossi et al. (1994) J Biol Chem 269:10444-10450; and Perez et al. (1992) J Cell Sci 102:717-722). The transcytosis polypeptide may also be a non-naturally-occurring membrane-translocating sequence (MTS), such as the peptide sequences disclosed in U.S. Patent No. 6,248,558.

In another embodiment, a polypeptide of the invention is labeled with an isotopic label to facilitate its detection and or structural characterization using nuclear magnetic resonance or another applicable technique. Exemplary isotopic labels include radioisotopic

labels such as, for example, potassium-40 (^{40}K), carbon-14 (^{14}C), tritium (^3H), sulphur-35 (^{35}S), phosphorus-32 (^{32}P), technetium-99m ($^{99\text{m}}\text{Tc}$), thallium-201 (^{201}Tl), gallium-67 (^{67}Ga), indium-111 (^{111}In), iodine-123 (^{123}I), iodine-131 (^{131}I), yttrium-90 (^{90}Y), samarium-153 (^{153}Sm), rhenium-186 (^{186}Re), rhenium-188 (^{188}Re), dysprosium-165 (^{165}Dy) and holmium-166 (^{166}Ho). The isotopic label may also be an atom with non zero nuclear spin, including, for example, hydrogen-1 (^1H), hydrogen-2 (^2H), hydrogen-3 (^3H), phosphorus-31 (^{31}P), sodium-23 (^{23}Na), nitrogen-14 (^{14}N), nitrogen-15 (^{15}N), carbon-13 (^{13}C) and fluorine-19 (^{19}F). In certain embodiments, the polypeptide is uniformly labeled with an isotopic label, for example, wherein at least 50%, 70%, 80%, 90%, 95%, or 98% of the possible labels in the polypeptide are labeled, e.g., wherein at least 50%, 70%, 80%, 90%, 95%, or 98% of the nitrogen atoms in the polypeptide are ^{15}N , and/or wherein at least 50%, 70%, 80%, 90%, 95%, or 98% of the carbon atoms in the polypeptide are ^{13}C , and/or wherein at least 50%, 70%, 80%, 90%, 95%, or 98% of the hydrogen atoms in the polypeptide are ^2H . In other embodiments, the isotopic label is located in one or more specific locations within the polypeptide, for example, the label may be specifically incorporated into one or more of the leucine residues of the polypeptide. The invention also encompasses the embodiment wherein a single polypeptide comprises two, three or more different isotopic labels, for example, the polypeptide comprises both ^{15}N and ^{13}C labeling.

In yet another embodiment, the polypeptides of the invention are labeled to facilitate structural characterization using x-ray crystallography or another applicable technique. Exemplary labels include heavy atom labels such as, for example, cobalt, selenium, krypton, bromine, strontium, molybdenum, ruthenium, rhodium, palladium, silver, cadmium, tin, iodine, xenon, barium, lanthanum, cerium, praseodymium, neodymium, samarium, europium, gadolinium, terbium, dysprosium, holmium, erbium, thulium, ytterbium, lutetium, tantalum, tungsten, rhenium, osmium, iridium, platinum, gold, mercury, thallium, lead, thorium and uranium. In an exemplary embodiment, the polypeptide is labeled with seleno-methionine.

A variety of methods are available for preparing a polypeptide with a label, such as a radioisotopic label or heavy atom label. For example, in one such method, an expression vector comprising a nucleic acid encoding a polypeptide is introduced into a host cell, and the host cell is cultured in a cell culture medium in the presence of a source of the label, thereby generating a labeled polypeptide. As indicated above, the extent to which a polypeptide may be labeled may vary.

In still another embodiment, the polypeptides of the invention are labeled with a fluorescent label to facilitate their detection, purification, or structural characterization. In an exemplary embodiment, a polypeptide of the invention is fused to a heterologous polypeptide sequence which produces a detectable fluorescent signal, including, for example, green fluorescent protein (GFP), enhanced green fluorescent protein (EGFP), Renilla Reniformis green fluorescent protein, GFPmut2, GFPuv4, enhanced yellow fluorescent protein (EYFP), enhanced cyan fluorescent protein (ECFP), enhanced blue fluorescent protein (EBFP), citrine and red fluorescent protein from discosoma (dsRED).

In other embodiments, the invention provides for polypeptides of the invention immobilized onto a solid surface, including, microtiter plates, slides, beads, films, etc. The polypeptides of the invention may be immobilized onto a "chip" as part of an array. An array, having a plurality of addresses, may comprise one or more polypeptides of the invention in one or more of those addresses. In one embodiment, the chip comprises one or more polypeptides of the invention as part of an array of *P. aeruginosa* polypeptide sequences.

In other embodiments, the invention provides for polypeptides of the invention immobilized onto a solid surface, including, plates, microtiter plates, slides, beads, particles, spheres, films, strands, precipitates, gels, sheets, tubing, containers, capillaries, pads, slices, etc. The polypeptides of the invention may be immobilized onto a "chip" as part of an array. An array, having a plurality of addresses, may comprise one or more polypeptides of the invention in one or more of those addresses. In one embodiment, the chip comprises one or more polypeptides of the invention as part of an array that contains at least some polypeptide sequences from *P. aeruginosa*.

In still other embodiments, the invention comprises the polypeptide sequences of the invention in computer readable format. The invention also encompasses a database comprising the polypeptide sequences of the invention.

In other embodiments, the invention relates to the polypeptides of the invention contained within a vessels useful for manipulation of the polypeptide sample. For example, the polypeptides of the invention may be contained within a microtiter plate to facilitate detection, screening or purification of the polypeptide. The polypeptides may also be contained within a syringe as a container suitable for administering the polypeptide to a subject in order to generate antibodies or as part of a vaccination regimen. The

polypeptides may also be contained within an NMR tube in order to enable characterization by nuclear magnetic resonance techniques.

In still other embodiments, the invention relates to a crystallized polypeptide of the invention and crystallized polypeptides which have been mounted for examination by x-ray crystallography as described further below. In certain instances, a polypeptide of the invention in crystal form may be single crystals of various dimensions (e.g., micro-crystals) or may be an aggregate of crystalline material. In another aspect, the present invention contemplates a crystallized complex including a polypeptide of the invention and one or more of the following: a co-factor (such as a salt, metal, nucleotide, oligonucleotide or polypeptide), a modulator, or a small molecule. In another aspect, the present invention contemplates a crystallized complex including a polypeptide of the invention and any other molecule or atom (such as a metal ion) that associates with the polypeptide *in vivo*.

In certain embodiments, polypeptides of the invention may be synthesized chemically, ribosomally in a cell free system, or ribosomally within a cell. Chemical synthesis of polypeptides of the invention may be carried out using a variety of art recognized methods, including stepwise solid phase synthesis, semi-synthesis through the conformationally-assisted re-ligation of peptide fragments, enzymatic ligation of cloned or synthetic peptide segments, and chemical ligation. Native chemical ligation employs a chemoselective reaction of two unprotected peptide segments to produce a transient thioester-linked intermediate. The transient thioester-linked intermediate then spontaneously undergoes a rearrangement to provide the full length ligation product having a native peptide bond at the ligation site. Full length ligation products are chemically identical to proteins produced by cell free synthesis. Full length ligation products may be refolded and/or oxidized, as allowed, to form native disulfide-containing protein molecules. (see e.g., U.S. Patent Nos. 6,184,344 and 6,174,530; and T. W. Muir et al., *Curr. Opin. Biotech.* (1993): vol. 4, p 420; M. Miller, et al., *Science* (1989): vol. 246, p 1149; A. Wlodawer, et al., *Science* (1989): vol. 245, p 616; L. H. Huang, et al., *Biochemistry* (1991): vol. 30, p 7402; M. Schnolzer, et al., *Int. J. Pept. Prot. Res.* (1992): vol. 40, p 180-193; K. Rajarathnam, et al., *Science* (1994): vol. 264, p 90; R. E. Offord, "Chemical Approaches to Protein Engineering", in *Protein Design and the Development of New therapeutics and Vaccines*, J. B. Hook, G. Poste, Eds., (Plenum Press, New York, 1990) pp. 253-282; C. J. A. Wallace, et al., *J. Biol. Chem.* (1992): vol. 267, p 3852; L. Abrahmsen, et al., *Biochemistry* (1991): vol. 30, p 4151; T. K. Chang, et al., *Proc. Natl. Acad. Sci. USA*

(1994) 91: 12544-12548; M. Schnlzer, et al., *Science* (1992): vol., 3256, p 221; and K. Akaji, et al., *Chem. Pharm. Bull. (Tokyo)* (1985) 33: 184).

In certain embodiments, it may be advantageous to provide naturally-occurring or experimentally-derived homologs of a polypeptide of the invention. Such homologs may function in a limited capacity as a modulator to promote or inhibit a subset of the biological activities of the naturally-occurring form of the polypeptide. Thus, specific biological effects may be elicited by treatment with a homolog of limited function, and with fewer side effects relative to treatment with agonists or antagonists which are directed to all of the biological activities of a polypeptide of the invention. For instance, antagonistic homologs may be generated which interfere with the ability of the wild-type polypeptide of the invention to associate with certain proteins, but which do not substantially interfere with the formation of complexes between the native polypeptide and other cellular proteins.

Another aspect of the invention relates to polypeptides derived from the full-length polypeptides of the invention. Isolated peptidyl portions of those polypeptides may be obtained by screening polypeptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such polypeptides. In addition, fragments may be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, proteins may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or may be divided into overlapping fragments of a desired length. The fragments may be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments having a desired property, for example, the capability of functioning as a modulator of the polypeptides of the invention. In an illustrative embodiment, peptidyl portions of a protein of the invention may be tested for binding activity, as well as inhibitory ability, by expression as, for example, thioredoxin fusion proteins, each of which contains a discrete fragment of a protein of the invention (see, for example, U.S. Patents 5,270,181 and 5,292,646; and PCT publication WO94/02502).

In another embodiment, truncated polypeptides may be prepared. Truncated polypeptides have from 1 to 20 or more amino acid residues removed from either or both the N- and C-termini. Such truncated polypeptides may prove more amenable to expression, purification or characterization than the full-length polypeptide. For example, truncated polypeptides may prove more amenable than the full-length polypeptide to crystallization, to yielding high quality diffracting crystals or to yielding an HSQC

spectrum with high intensity peaks and minimally overlapping peaks. In addition, the use of truncated polypeptides may also identify stable and active domains of the full-length polypeptide that may be more amenable to characterization.

It is also possible to modify the structure of the polypeptides of the invention for such purposes as enhancing therapeutic or prophylactic efficacy, or stability (e.g., *ex vivo* shelf life, resistance to proteolytic degradation *in vivo*, etc.). Such modified polypeptides, when designed to retain at least one activity of the naturally-occurring form of the protein, are considered "functional equivalents" of the polypeptides described in more detail herein. Such modified polypeptides may be produced, for instance, by amino acid substitution, deletion, or addition, which substitutions may consist in whole or part by conservative amino acid substitutions.

For instance, it is reasonable to expect that an isolated conservative amino acid substitution, such as replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, will not have a major affect on the biological activity of the resulting molecule. Whether a change in the amino acid sequence of a polypeptide results in a functional homolog may be readily determined by assessing the ability of the variant polypeptide to produce a response similar to that of the wild-type protein. Polypeptides in which more than one replacement has taken place may readily be tested in the same manner.

This invention further contemplates a method of generating sets of combinatorial mutants of polypeptides of the invention, as well as truncation mutants, and is especially useful for identifying potential variant sequences (e.g. homologs). The purpose of screening such combinatorial libraries is to generate, for example, homologs which may modulate the activity of a polypeptide of the invention, or alternatively, which possess novel activities altogether. Combinatorially-derived homologs may be generated which have a selective potency relative to a naturally-occurring protein. Such homologs may be used in the development of therapeutics.

Likewise, mutagenesis may give rise to homologs which have intracellular half-lives dramatically different than the corresponding wild-type protein. For example, the altered protein may be rendered either more stable or less stable to proteolytic degradation or other cellular process which result in destruction of, or otherwise inactivation of the protein. Such homologs, and the genes which encode them, may be utilized to alter protein

expression by modulating the half-life of the protein. As above, such proteins may be used for the development of therapeutics or treatment.

In similar fashion, protein homologs may be generated by the present combinatorial approach to act as antagonists, in that they are able to interfere with the activity of the corresponding wild-type protein.

In a representative embodiment of this method, the amino acid sequences for a population of protein homologs are aligned, preferably to promote the highest homology possible. Such a population of variants may include, for example, homologs from one or more species, or homologs from the same species but which differ due to mutation. Amino acids which appear at each position of the aligned sequences are selected to create a degenerate set of combinatorial sequences. In certain embodiments, the combinatorial library is produced by way of a degenerate library of genes encoding a library of polypeptides which each include at least a portion of potential protein sequences. For instance, a mixture of synthetic oligonucleotides may be enzymatically ligated into gene sequences such that the degenerate set of potential nucleotide sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g. for phage display).

There are many ways by which the library of potential homologs may be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence may be carried out in an automatic DNA synthesizer, and the synthetic genes may then be ligated into an appropriate vector for expression. One purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential protein sequences. The synthesis of degenerate oligonucleotides is well known in the art (see for example, Narang, SA (1983) *Tetrahedron* 39:3; Itakura et al., (1981) *Recombinant DNA, Proc. 3rd Cleveland Sympos. Macromolecules*, ed. AG Walton, Amsterdam: Elsevier pp. 273-289; Itakura et al., (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al., (1984) *Science* 198:1056; Ike et al., (1983) *Nucleic Acid Res.* 11:477). Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al., (1990) *Science* 249:386-390; Roberts et al., (1992) *PNAS USA* 89:2429-2433; Devlin et al., (1990) *Science* 249: 404-406; Cwirla et al., (1990) *PNAS USA* 87: 6378-6382; as well as U.S. Patent Nos: 5,223,409, 5,198,346, and 5,096,815).

Alternatively, other forms of mutagenesis may be utilized to generate a combinatorial library. For example, protein homologs (both agonist and antagonist forms) may be generated and isolated from a library by screening using, for example, alanine scanning mutagenesis and the like (Ruf et al., (1994) *Biochemistry* 33:1565-1572; Wang et al., (1994) *J. Biol. Chem.* 269:3095-3099; Balint et al., (1993) *Gene* 137:109-118; Grodberg et al., (1993) *Eur. J. Biochem.* 218:597-601; Nagashima et al., (1993) *J. Biol. Chem.* 268:2888-2892; Lowman et al., (1991) *Biochemistry* 30:10832-10838; and Cunningham et al., (1989) *Science* 244:1081-1085), by linker scanning mutagenesis (Gustin et al., (1993) *Virology* 193:653-660; Brown et al., (1992) *Mol. Cell Biol.* 12:2644-2652; McKnight et al., (1982) *Science* 232:316); by saturation mutagenesis (Meyers et al., (1986) *Science* 232:613); by PCR mutagenesis (Leung et al., (1989) *Method Cell Mol Biol* 1:11-19); or by random mutagenesis (Miller et al., (1992) *A Short Course in Bacterial Genetics*, CSHL Press, Cold Spring Harbor, NY; and Greener et al., (1994) *Strategies in Mol Biol* 7:32-34). Linker scanning mutagenesis, particularly in a combinatorial setting, is an attractive method for identifying truncated (bioactive) forms of proteins.

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations and truncations, and for screening cDNA libraries for gene products having a certain property. Such techniques will be generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of protein homologs. The most widely used techniques for screening large gene libraries typically comprises cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays described below are amenable to high throughput analysis as necessary to screen large numbers of degenerate sequences created by combinatorial mutagenesis techniques.

In an illustrative embodiment of a screening assay, candidate combinatorial gene products are displayed on the surface of a cell and the ability of particular cells or viral particles to bind to the combinatorial gene product is detected in a "panning assay". For instance, the gene library may be cloned into the gene for a surface membrane protein of a bacterial cell (Ladner et al., WO 88/06630; Fuchs et al., (1991) *Bio/Technology* 9:1370-1371; and Goward et al., (1992) *TIBS* 18:136-140), and the resulting fusion protein detected

by panning, e.g. using a fluorescently labeled molecule which binds the cell surface protein, e.g. FITC-substrate, to score for potentially functional homologs. Cells may be visually inspected and separated under a fluorescence microscope, or, when the morphology of the cell permits, separated by a fluorescence-activated cell sorter. This method may be used to identify substrates or other polypeptides that can interact with a polypeptide of the invention.

In similar fashion, the gene library may be expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences may be expressed on the surface of infectious phage, thereby conferring two benefits. First, because these phage may be applied to affinity matrices at very high concentrations, a large number of phage may be screened at one time. Second, because each infectious phage displays the combinatorial gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage may be amplified by another round of infection. The group of almost identical *E. coli* filamentous phages M13, fd, and f1 are most often used in phage display libraries, as either of the phage gIII or gVIII coat proteins may be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle (Ladner et al., PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al., (1992) *J. Biol. Chem.* 267:16007-16010; Griffiths et al., (1993) *EMBO J.* 12:725-734; Clackson et al., (1991) *Nature* 352:624-628; and Barbas et al., (1992) *PNAS USA* 89:4457-4461). Other phage coat proteins may be used as appropriate.

The invention also provides for reduction of the polypeptides of the invention to generate mimetics, e.g. peptide or non-peptide agents, which are able to mimic binding of the authentic protein to another cellular partner. Such mutagenic techniques as described above, as well as the thioredoxin system, are also particularly useful for mapping the determinants of a protein which participates in a protein-protein interaction with another protein. To illustrate, the critical residues of a protein which are involved in molecular recognition of a substrate protein may be determined and used to generate peptidomimetics that may bind to the substrate protein. The peptidomimetic may then be used as an inhibitor of the wild-type protein by binding to the substrate and covering up the critical residues needed for interaction with the wild-type protein, thereby preventing interaction of the protein and the substrate. By employing, for example, scanning mutagenesis to map the amino acid residues of a protein which are involved in binding a substrate polypeptide,

peptidomimetic compounds may be generated which mimic those residues in binding to the substrate. For instance, non-hydrolyzable peptide analogs of such residues may be generated using benzodiazepine (e.g., see Freidinger et al., in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al., in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gamma lactam rings (Garvey et al., in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al., (1986) *J. Med. Chem.* 29:295; and Ewenson et al., in *Peptides: Structure and Function* (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, IL, 1985), β -turn dipeptide cores (Nagai et al., (1985) *Tetrahedron Lett* 26:647; and Sato et al., (1986) *J Chem Soc Perkin Trans* 1:1231), and β -aminoalcohols (Gordon et al., (1985) *Biochem Biophys Res Commun* 126:419; and Dann et al., (1986) *Biochem Biophys Res Commun* 134:71).

The activity of a polypeptide of the invention may be identified and/or assayed using a variety of methods well known to the skilled artisan. For example, information about the activity of non-essential genes may be assayed by creating a null mutant strain of bacteria expressing a mutant form of, or lacking expression of, a protein of interest. The resulting phenotype of the null mutant strain may provide information about the activity of the mutated gene product. Essential genes may be studied by creating a bacterial strain with a conditional mutation in the gene of interest. The bacterial strain may be grown under permissive and non-permissive conditions and the change in phenotype under the non-permissive conditions may be used to identify and/or assay the activity of the gene product.

In an alternative embodiment, the activity of a protein may be assayed using an appropriate substrate or binding partner or other reagent suitable to test for the suspected activity. For catalytic activity, the assay is typically designed so that the enzymatic reaction produces a detectable signal. For example, mixture of a kinase with a substrate in the presence of ^{32}P will result in incorporation of the ^{32}P into the substrate. The labeled substrate may then be separated from the free ^{32}P and the presence and/or amount of radiolabeled substrate may be detected using a scintillation counter or a phosphorimager. Similar assays may be designed to identify and/or assay the activity of a wide variety of enzymatic activities. Based on the teachings herein, the skilled artisan would readily be able to develop an appropriate assay for a polypeptide of the invention.

In another embodiment, the activity of a polypeptide of the invention may be determined by assaying for the level of expression of RNA and/or protein molecules. Transcription levels may be determined, for example, using Northern blots, hybridization to an oligonucleotide array or by assaying for the level of a resulting protein product. Translation levels may be determined, for example, using Western blotting or by identifying a detectable signal produced by a protein product (e.g., fluorescence, luminescence, enzymatic activity, etc.). Depending on the particular situation, it may be desirable to detect the level of transcription and/or translation of a single gene or of multiple genes.

Alternatively, it may be desirable to measure the overall rate of DNA replication, transcription and/or translation in a cell. In general this may be accomplished by growing the cell in the presence of a detectable metabolite which is incorporated into the resultant DNA, RNA, or protein product. For example, the rate of DNA synthesis may be determined by growing cells in the presence of BrdU which is incorporated into the newly synthesized DNA. The amount of BrdU may then be determined histochemically using an anti-BrdU antibody.

In general, the biological activity of a polypeptide encoded by SEQ ID NO: 2 or SEQ ID NO: 4, and possibly other polypeptides of the invention, is triosephosphate isomerase, having the gene designation of *tpiA*. The polypeptide encoded by SEQ ID NO: 2 or SEQ ID NO: 4, and possibly other polypeptides of the invention, may be further characterized as being part of the COG category "carbohydrate transport and metabolism", with COG ID No. COG0149. The foregoing annotations were determined in accordance with the procedure described in EXAMPLE 16. This functionality assignment has been confirmed by completion of the X-ray structure of a polypeptide of the invention, as described in more detail below. In one aspect, the present invention contemplates a polypeptide having biological activity, or is a component of a protein complex having biological activity, substantially similar to or identical to triosephosphate isomerase. Alternatively, the polypeptide catalyzes, or is a component of a protein complex that catalyzes, a reaction that is substantially the same type of, or is the same as, the reaction catalyzed by triosephosphate isomerase. Other biological activities of polypeptides of the invention are described herein, or will be reasonably apparent to those skilled in the art in light of the present disclosure.

Glycolysis, in addition to its role as a important biological route for the metabolism of hexoses, provides the cell with intermediates of central metabolism for the synthesis of amino acids, vitamins, nucleotides, and cell wall constituents. Triosephosphate isomerase (TPI, EC 5.3.1.1) catalyzes interconversion of glyceraldehydes-3-phosphate and dihydroxyacetone phosphate, both formed from catabolism of fructose-1,6-diphosphate. The main metabolic role for this enzyme is to convert essentially nonmetabolizing dihydroxyacetonephosphate into glyceraldehyde-3-phosphate. By playing a critical role, such an enzyme presents a favorable target for therapeutics and diagnostics.

3. Nucleic Acids of the Invention

One aspect of the invention pertains to isolated nucleic acids of the invention. For example, the present invention contemplates an isolated nucleic acid comprising (a) the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, (b) a nucleotide sequence at least 80% identical to SEQ ID NO: 1 or SEQ ID NO: 3, (c) a nucleotide sequence that hybridizes under stringent conditions to SEQ ID NO: 1 or SEQ ID NO: 3, or (d) the complement of the nucleotide sequence of (a), (b) or (c). In certain embodiments, nucleic acids of the invention may be labeled, with for example, a radioactive, chemiluminescent or fluorescent label.

It may be that case that the nucleic acid sequence of SEQ ID NO: 3 differs from that of SEQ ID NO: 1 by one or more nucleic acid residues. SEQ ID NO: 3 is determined experimentally, and SEQ ID NO: 1 obtained as described in EXAMPLE 1. In such a case, the present invention contemplates the specific nucleic acid sequences of SEQ ID NO: 1 and SEQ ID NO: 3, and variants thereof, as well as any differences in the applicable amino acid sequences encoded thereby.

In another aspect, the present invention contemplates an isolated nucleic acid that specifically hybridizes under stringent conditions to at least ten nucleotides of SEQ ID NO: 1 or SEQ ID NO: 3, or the complement thereof, which nucleic acid can specifically detect or amplify SEQ ID NO: 1 or SEQ ID NO: 3, or the complement thereof. In yet another aspect, the present invention contemplates such an isolated nucleic acid comprising a nucleotide sequence encoding a fragment of SEQ ID NO: 2 or SEQ ID NO: 4 at least 8 residues in length. The present invention further contemplates a method of hybridizing an oligonucleotide with a nucleic acid of the invention comprising: (a) providing a single-

stranded oligonucleotide at least eight nucleotides in length, the oligonucleotide being complementary to a portion of a nucleic acid of the invention; and (b) contacting the oligonucleotide with a sample comprising a nucleic acid of the acid under conditions that permit hybridization of the oligonucleotide with the nucleic acid of the invention.

Isolated nucleic acids which differ from the nucleic acids of the invention due to degeneracy in the genetic code are also within the scope of the invention. For example, a number of amino acids are designated by more than one triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and CAC are synonyms for histidine) may result in "silent" mutations which do not affect the amino acid sequence of the protein. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the polypeptides of the invention will exist. One skilled in the art will appreciate that these variations in one or more nucleotides (from less than 1% up to about 3 or 5% or possibly more of the nucleotides) of the nucleic acids encoding a particular protein of the invention may exist among a given species due to natural allelic variation. Any and all such nucleotide variations and resulting amino acid polymorphisms are within the scope of this invention.

Bias in codon choice within genes in a single species appears related to the level of expression of the protein encoded by that gene. Accordingly, the invention encompasses nucleic acid sequences which have been optimized for improved expression in a host cell by altering the frequency of codon usage in the nucleic acid sequence to approach the frequency of preferred codon usage of the host cell. Due to codon degeneracy, it is possible to optimize the nucleotide sequence without affecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleotide sequence that encodes all or a substantial portion of the amino acid sequence set forth in SEQ ID NO: 2, SEQ ID NO: 4 or other polypeptides of the invention.

The present invention pertains to nucleic acids encoding proteins derived from *P. aeruginosa* and which have amino acid sequences evolutionarily related to a polypeptide of the invention, wherein "evolutionarily related to", refers to proteins having different amino acid sequences which have arisen naturally (e.g. by allelic variance or by differential splicing), as well as mutational variants of the proteins of the invention which are derived, for example, by combinatorial mutagenesis.

Fragments of the polynucleotides of the invention encoding a biologically active portion of the subject polypeptides are also within the scope of the invention. As used herein, a fragment of a nucleic acid of the invention encoding an active portion of a polypeptide of the invention refers to a nucleotide sequence having fewer nucleotides than the nucleotide sequence encoding the full length amino acid sequence of a polypeptide of the invention, for example, SEQ ID NO: 2 or SEQ ID NO: 4, and which encodes a polypeptide which retains at least a portion of a biological activity of the full-length protein as defined herein, or alternatively, which is functional as a modulator of the biological activity of the full-length protein. For example, such fragments include a polypeptide containing a domain of the full-length protein from which the polypeptide is derived that mediates the interaction of the protein with another molecule (e.g., polypeptide, DNA, RNA, etc.). In another embodiment, the present invention contemplates an isolated nucleic acid that encodes a polypeptide having a biological activity of a protein having the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4, or alternatively biological activity of triosephosphate isomerase.

Nucleic acids within the scope of the invention may also contain linker sequences, modified restriction endonuclease sites and other sequences useful for molecular cloning, expression or purification of such recombinant polypeptides.

A nucleic acid encoding a polypeptide of the invention may be obtained from mRNA or genomic DNA from any organism in accordance with protocols described herein, as well as those generally known to those skilled in the art. A cDNA encoding a polypeptide of the invention, for example, may be obtained by isolating total mRNA from an organism, e.g. a bacteria, virus, mammal, etc. Double stranded cDNAs may then be prepared from the total mRNA, and subsequently inserted into a suitable plasmid or bacteriophage vector using any one of a number of known techniques. A gene encoding a polypeptide of the invention may also be cloned using established polymerase chain reaction techniques in accordance with the nucleotide sequence information provided by the invention. In one aspect, the present invention contemplates a method for amplification of a nucleic acid of the invention, or a fragment thereof, comprising: (a) providing a pair of single stranded oligonucleotides, each of which is at least eight nucleotides in length, complementary to sequences of a nucleic acid of the invention, and wherein the sequences to which the oligonucleotides are complementary are at least ten nucleotides apart; and (b) contacting the oligonucleotides with a sample comprising a nucleic acid comprising the

nucleic acid of the invention under conditions which permit amplification of the region located between the pair of oligonucleotides, thereby amplifying the nucleic acid.

Another aspect of the invention relates to the use of nucleic acids of the invention in "antisense therapy". As used herein, antisense therapy refers to administration or *in situ* generation of oligonucleotide probes or their derivatives which specifically hybridize or otherwise bind under cellular conditions with the cellular mRNA and/or genomic DNA encoding one of the polypeptides of the invention so as to inhibit expression of that polypeptide, e.g. by inhibiting transcription and/or translation. The binding may be by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix. In general, antisense therapy refers to the range of techniques generally employed in the art, and includes any therapy which relies on specific binding to oligonucleotide sequences.

An antisense construct of the present invention may be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the mRNA which encodes a polypeptide of the invention. Alternatively, the antisense construct may be an oligonucleotide probe which is generated *ex vivo* and which, when introduced into the cell causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences encoding a polypeptide of the invention. Such oligonucleotide probes may be modified oligonucleotides which are resistant to endogenous nucleases, e.g. exonucleases and/or endonucleases, and are therefore stable *in vivo*. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to constructing oligomers useful in antisense therapy have been reviewed, for example, by van der Krol et al., (1988) *Biotechniques* 6:958-976; and Stein et al., (1988) *Cancer Res* 48:2659-2668.

In a further aspect, the invention provides double stranded small interfering RNAs (siRNAs), and methods for administering the same. siRNAs decrease or block gene expression. While not wishing to be bound by theory, it is generally thought that siRNAs inhibit gene expression by mediating sequence specific mRNA degradation. RNA interference (RNAi) is the process of sequence-specific, post-transcriptional gene silencing, particularly in animals and plants, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene (Elbashir et al. *Nature* 2001; 411(6836): 494-

8). Accordingly, it is understood that siRNAs and long dsRNAs having substantial sequence identity to all or a portion of SEQ ID NO: 1 or SEQ ID NO: 3 may be used to inhibit the expression of a nucleic acid of the invention, and particularly when the polynucleotide is expressed in a mammalian or plant cell.

The nucleic acids of the invention may be used as diagnostic reagents to detect the presence or absence of the target DNA or RNA sequences to which they specifically bind, such as for determining the level of expression of a nucleic acid of the invention. In one aspect, the present invention contemplates a method for detecting the presence of a nucleic acid of the invention or a portion thereof in a sample, the method comprising: (a) providing an oligonucleotide at least eight nucleotides in length, the oligonucleotide being complementary to a portion of a nucleic acid of the invention; (b) contacting the oligonucleotide with a sample comprising at least one nucleic acid under conditions that permit hybridization of the oligonucleotide with a nucleic acid comprising a nucleotide sequence complementary thereto; and (c) detecting hybridization of the oligonucleotide to a nucleic acid in the sample, thereby detecting the presence of a nucleic acid of the invention or a portion thereof in the sample. In another aspect, the present invention contemplates a method for detecting the presence of a nucleic acid of the invention or a portion thereof in a sample, the method comprising: (a) providing a pair of single stranded oligonucleotides, each of which is at least eight nucleotides in length, complementary to sequences of a nucleic acid of the invention, and wherein the sequences to which the oligonucleotides are complementary are at least ten nucleotides apart; and (b) contacting the oligonucleotides with a sample comprising at least one nucleic acid under hybridization conditions; (c) amplifying the nucleotide sequence between the two oligonucleotide primers; and (d) detecting the presence of the amplified sequence, thereby detecting the presence of a nucleic acid comprising the nucleic acid of the invention or a portion thereof in the sample.

In another aspect of the invention, the subject nucleic acid is provided in an expression vector comprising a nucleotide sequence encoding a polypeptide of the invention and operably linked to at least one regulatory sequence. It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed and/or the type of protein desired to be expressed. The vector's copy number, the ability to control that copy number and the expression of any other protein encoded by the vector, such as antibiotic markers, should be considered.

The subject nucleic acids may be used to cause expression and over-expression of a polypeptide of the invention in cells propagated in culture, e.g. to produce proteins or polypeptides, including fusion proteins or polypeptides.

This invention pertains to a host cell transfected with a recombinant gene in order to express a polypeptide of the invention. The host cell may be any prokaryotic or eukaryotic cell. For example, a polypeptide of the invention may be expressed in bacterial cells, such as *E. coli*, insect cells (baculovirus), yeast, or mammalian cells. In those instances when the host cell is human, it may or may not be in a live subject. Other suitable host cells are known to those skilled in the art. Additionally, the host cell may be supplemented with tRNA molecules not typically found in the host so as to optimize expression of the polypeptide. Other methods suitable for maximizing expression of the polypeptide will be known to those in the art.

The present invention further pertains to methods of producing the polypeptides of the invention. For example, a host cell transfected with an expression vector encoding a polypeptide of the invention may be cultured under appropriate conditions to allow expression of the polypeptide to occur. The polypeptide may be secreted and isolated from a mixture of cells and medium containing the polypeptide. Alternatively, the polypeptide may be retained cytoplasmically and the cells harvested, lysed and the protein isolated.

A cell culture includes host cells, media and other byproducts. Suitable media for cell culture are well known in the art. The polypeptide may be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins, including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for particular epitopes of a polypeptide of the invention.

Thus, a nucleotide sequence encoding all or a selected portion of polypeptide of the invention, may be used to produce a recombinant form of the protein via microbial or eukaryotic cellular processes. Ligating the sequence into a polynucleotide construct, such as an expression vector, and transforming or transfecting into hosts, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial cells), are standard procedures. Similar procedures, or modifications thereof, may be employed to prepare recombinant polypeptides of the invention by microbial means or tissue-culture technology.

Expression vehicles for production of a recombinant protein include plasmids and other vectors. For instance, suitable vectors for the expression of a polypeptide of the invention include plasmids of the types: pBR322-derived plasmids, pEMBL-derived plasmids, pEX-derived plasmids, pBTac-derived plasmids and pUC-derived plasmids for expression in prokaryotic cells, such as *E. coli*.

A number of vectors exist for the expression of recombinant proteins in yeast. For instance, YEP24, YIP5, YEP51, YEP52, pYES2, and YRP17 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae* (see, for example, Broach et al., (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye Academic Press, p. 83). These vectors may replicate in *E. coli* due the presence of the pBR322 ori, and in *S. cerevisiae* due to the replication determinant of the yeast 2 micron plasmid. In addition, drug resistance markers such as ampicillin may be used.

In certain embodiments, mammalian expression vectors contain both prokaryotic sequences to facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The pcDNAI/amp, pcDNAI/neo, pRc/CMV, pSV2gpt, pSV2neo, pSV2-dhfr, pTk2, pRSVneo, pMSG, pSVT7, pko-neo and pHyg derived vectors are examples of mammalian expression vectors suitable for transfection of eukaryotic cells. Some of these vectors are modified with sequences from bacterial plasmids, such as pBR322, to facilitate replication and drug resistance selection in both prokaryotic and eukaryotic cells. Alternatively, derivatives of viruses such as the bovine papilloma virus (BPV-1), or Epstein-Barr virus (pHEBo, pREP-derived and p205) can be used for transient expression of proteins in eukaryotic cells. The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures, see *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press, 1989) Chapters 16 and 17. In some instances, it may be desirable to express the recombinant protein by the use of a baculovirus expression system. Examples of such baculovirus expression systems include pVL-derived vectors (such as pVL1392, pVL1393 and pVL941), pAcUW-derived vectors (such as pAcUW1), and pBlueBac-derived vectors (such as the β -gal containing pBlueBac III).

In another variation, protein production may be achieved using *in vitro* translation systems. *In vitro* translation systems are, generally, a translation system which is a cell-free

extract containing at least the minimum elements necessary for translation of an RNA molecule into a protein. An *in vitro* translation system typically comprises at least ribosomes, tRNAs, initiator methionyl-tRNA^{Met}, proteins or complexes involved in translation, e.g., eIF2, eIF3, the cap-binding (CB) complex, comprising the cap-binding protein (CBP) and eukaryotic initiation factor 4F (eIF4F). A variety of *in vitro* translation systems are well known in the art and include commercially available kits. Examples of *in vitro* translation systems include eukaryotic lysates, such as rabbit reticulocyte lysates, rabbit oocyte lysates, human cell lysates, insect cell lysates and wheat germ extracts. Lysates are commercially available from manufacturers such as Promega Corp., Madison, Wis.; Stratagene, La Jolla, Calif.; Amersham, Arlington Heights, Ill.; and GIBCO/BRL, Grand Island, N.Y. *In vitro* translation systems typically comprise macromolecules, such as enzymes, translation, initiation and elongation factors, chemical reagents, and ribosomes. In addition, an *in vitro* transcription system may be used. Such systems typically comprise at least an RNA polymerase holoenzyme, ribonucleotides and any necessary transcription initiation, elongation and termination factors. *In vitro* transcription and translation may be coupled in a one-pot reaction to produce proteins from one or more isolated DNAs.

When expression of a carboxy terminal fragment of a polypeptide is desired, i.e. a truncation mutant, it may be necessary to add a start codon (ATG) to the oligonucleotide fragment containing the desired sequence to be expressed. It is well known in the art that a methionine at the N-terminal position may be enzymatically cleaved by the use of the enzyme methionine aminopeptidase (MAP). MAP has been cloned from *E. coli* (Ben-Bassat et al., (1987) *J. Bacteriol.* 169:751-757) and *Salmonella typhimurium* and its *in vitro* activity has been demonstrated on recombinant proteins (Miller et al., (1987) *PNAS USA* 84:2718-1722). Therefore, removal of an N-terminal methionine, if desired, may be achieved either *in vivo* by expressing such recombinant polypeptides in a host which produces MAP (e.g., *E. coli* or CM89 or *S. cerevisiae*), or *in vitro* by use of purified MAP (e.g., procedure of Miller et al.).

Coding sequences for a polypeptide of interest may be incorporated as a part of a fusion gene including a nucleotide sequence encoding a different polypeptide. The present invention contemplates an isolated nucleic acid comprising a nucleic acid of the invention and at least one heterologous sequence encoding a heterologous peptide linked in frame to the nucleotide sequence of the nucleic acid of the invention so as to encode a fusion protein comprising the heterologous polypeptide. The heterologous polypeptide may be fused to

(a) the C-terminus of the polypeptide encoded by the nucleic acid of the invention, (b) the N-terminus of the polypeptide, or (c) the C-terminus and the N-terminus of the polypeptide. In certain instances, the heterologous sequence encodes a polypeptide permitting the detection, isolation, solubilization and/or stabilization of the polypeptide to which it is fused. In still other embodiments, the heterologous sequence encodes a polypeptide selected from the group consisting of a polyHis tag, myc, HA, GST, protein A, protein G, calmodulin-binding peptide, thioredoxin, maltose-binding protein, poly arginine, poly His-Asp, FLAG, a portion of an immunoglobulin protein, and a transcytosis peptide.

Fusion expression systems can be useful when it is desirable to produce an immunogenic fragment of a polypeptide of the invention. For example, the VP6 capsid protein of rotavirus may be used as an immunologic carrier protein for portions of polypeptide, either in the monomeric form or in the form of a viral particle. The nucleic acid sequences corresponding to the portion of a polypeptide of the invention to which antibodies are to be raised may be incorporated into a fusion gene construct which includes coding sequences for a late vaccinia virus structural protein to produce a set of recombinant viruses expressing fusion proteins comprising a portion of the protein as part of the virion. The Hepatitis B surface antigen may also be utilized in this role as well. Similarly, chimeric constructs coding for fusion proteins containing a portion of a polypeptide of the invention and the poliovirus capsid protein may be created to enhance immunogenicity (see, for example, EP Publication NO: 0259149; and Evans et al., (1989) *Nature* 339:385; Huang et al., (1988) *J. Virol.* 62:3855; and Schlienger et al., (1992) *J. Virol.* 66:2).

Fusion proteins may facilitate the expression and/or purification of proteins. For example, a polypeptide of the invention may be generated as a glutathione-S-transferase (GST) fusion protein. Such GST fusion proteins may be used to simplify purification of a polypeptide of the invention, such as through the use of glutathione-derivatized matrices (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al., (N.Y.: John Wiley & Sons, 1991)). In another embodiment, a fusion gene coding for a purification leader sequence, such as a poly-(His)/enterokinase cleavage site sequence at the N-terminus of the desired portion of the recombinant protein, may allow purification of the expressed fusion protein by affinity chromatography using a Ni^{2+} metal resin. The purification leader sequence may then be subsequently removed by treatment with enterokinase to provide the purified protein (e.g., see Hochuli et al., (1987) *J. Chromatography* 411: 177; and Janknecht et al., *PNAS USA* 88:8972).

Techniques for making fusion genes are well known. Essentially, the joining of various DNA fragments coding for different polypeptide sequences is performed in accordance with conventional techniques, employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene may be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments may be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which may subsequently be annealed to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al., John Wiley & Sons: 1992).

The present invention further contemplates a transgenic non-human animal having cells which harbor a transgene comprising a nucleic acid of the invention.

In other embodiments, the invention provides for nucleic acids of the invention immobilized onto a solid surface, including, plates, microtiter plates, slides, beads, particles, spheres, films, strands, precipitates, gels, sheets, tubing, containers, capillaries, pads, slices, etc. The nucleic acids of the invention may be immobilized onto a chip as part of an array. The array may comprise one or more polynucleotides of the invention as described herein. In one embodiment, the chip comprises one or more polynucleotides of the invention as part of an array of *P. aeruginosa* polynucleotide sequences.

In still other embodiments, the invention comprises the sequence of a nucleic acid of the invention in computer readable format. The invention also encompasses a database comprising the sequence of a nucleic acid of the invention.

4. Homology Searching of Nucleotide and Polypeptide Sequences

The nucleotide or amino acid sequences of the invention may be used as query sequences against databases such as GenBank, SwissProt, PDB, BLOCKS, and Pima II. These databases contain previously identified and annotated sequences that may be searched for regions of homology (similarity) using BLAST, which stands for Basic Local Alignment Search Tool (Altschul S F (1993) J Mol Evol 36:290-300; Altschul, S F et al (1990) J Mol Biol 215:403-10).

BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs which may be of prokaryotic (bacterial) or eukaryotic (animal, fungal or plant) origin. Other algorithms such as the one described in Smith, R. F. and T. F. Smith (1992; Protein Engineering 5:35-51) may be used when dealing with primary sequence patterns and secondary structure gap penalties. In the usual course using BLAST, sequences have lengths of at least 49 nucleotides and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

The BLAST approach, as detailed in Karlin and Altschul (1993; Proc Nat Acad Sci 90:5873-7) searches matches between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The threshold is typically set at about 10-25 for nucleotides and about 3-15 for peptides.

5. Analysis of Protein Properties

(a) Analysis of Proteins by Mass Spectrometry

Typically, protein characterization by mass spectroscopy first requires protein isolation followed by either chemical or enzymatic digestion of the protein into smaller peptide fragments, whereupon the peptide fragments may be analyzed by mass spectrometry to obtain a peptide map. Mass spectrometry may also be used to identify post-translational modifications (e.g., phosphorylation, etc.) of a polypeptide.

Various mass spectrometers may be used within the present invention. Representative examples include: triple quadrupole mass spectrometers, magnetic sector instruments (magnetic tandem mass spectrometer, JEOL, Peabody, Mass), ionspray mass spectrometers (Bruins et al., Anal. Chem. 59:2642-2647, 1987), electrospray mass spectrometers (including tandem, nano- and nano-electrospray tandem) (Fenn et al., Science 246:64-71, 1989), laser desorption time-of-flight mass spectrometers (Karas and Hillenkamp, Anal. Chem. 60:2299-2301, 1988), and a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (Extrel Corp., Pittsburgh, Mass.).

MALDI ionization is a technique in which samples of interest, in this case peptides and proteins, are co-crystallized with an acidified matrix. The matrix is typically a small molecule that absorbs at a specific wavelength, generally in the ultraviolet (UV) range, and dissipates the absorbed energy thermally. Typically a pulsed laser beam is used to transfer energy rapidly (i.e., a few ns) to the matrix. This transfer of energy causes the matrix to rapidly dissociate from the MALDI plate surface and results in a plume of matrix and the co-crystallized analytes being transferred into the gas phase. MALDI is considered a "soft-ionization" method that typically results in singly-charged species in the gas phase, most often resulting from a protonation reaction with the matrix. MALDI may be coupled in-line with time of flight (TOF) mass spectrometers. TOF detectors are based on the principle that an analyte moves with a velocity proportional to its mass. Analytes of higher mass move slower than analytes of lower mass and thus reach the detector later than lighter analytes. The present invention contemplates a composition comprising a polypeptide of the invention and a matrix suitable for mass spectrometry. In certain instances, the matrix is a nicotinic acid derivative or a cinnamic acid derivative.

MALDI-TOF MS is easily performed with modern mass spectrometers. Typically the samples of interest, in this case peptides or proteins, are mixed with a matrix and spotted onto a polished stainless steel plate (MALDI plate). Commercially available MALDI plates can presently hold up to 1536 samples per plate. Once spotted with sample, the MALDI sample plate is then introduced into the vacuum chamber of a MALDI mass spectrometer. The pulsed laser is then activated and the mass to charge ratios of the analytes are measured utilizing a time of flight detector. A mass spectrum representing the mass to charge ratios of the peptides/proteins is generated.

As mentioned above, MALDI can be utilized to measure the mass to charge ratios of both proteins and peptides. In the case of proteins, a mixture of intact protein and matrix are co-crystallized on a MALDI target (Karas, M. and Hillenkamp, F. Anal. Chem. 1988, 60 (20) 2299-2301). The spectrum resulting from this analysis is employed to determine the molecular weight of a whole protein. This molecular weight can then be compared to the theoretical weight of the protein and utilized in characterizing the analyte of interest, such as whether or not the protein has undergone post-translational modifications (e.g., example phosphorylation).

In certain embodiments, MALDI mass spectrometry is used for determination of peptide maps of digested proteins. The peptide masses are measured accurately using a

MALDI-TOF or a MALDI-Q-Star mass spectrometer, with detection precision down to the low ppm (parts per million) level. The ensemble of the peptide masses observed in a protein digest, such as a tryptic digest, may be used to search protein/DNA databases in a method called peptide mass fingerprinting. In this approach, protein entries in a database are ranked according to the number of experimental peptide masses that match the predicted trypsin digestion pattern. Commercially available software utilizes a search algorithm that provides a scoring scheme based on the size of the databases, the number of matching peptides, and the different peptides. Depending on the number of peptides observed, the accuracy of the measurement, and the size of the genome of the particular species, unambiguous protein identification may be obtained.

Statistical analysis may be performed upon each protein match to determine the validity of the match. Typical constraints include error tolerances within 0.1 Da for monoisotopic peptide masses, cysteines may be alkylated and searched as carboxyamidomethyl modifications, 0 or 1 missed enzyme cleavages, and no methionine oxidations allowed. Identified proteins may be stored automatically in a relational database with software links to SDS-PAGE images and ligand sequences. Often even a partial peptide map is specific enough for identification of the protein. If no protein match is found, a more error-tolerant search can be used, for example using fewer peptides or allowing a larger margin error with respect to mass accuracy.

Other mass spectroscopy methods such as tandem mass spectrometry or post source decay may be used to obtain sequence information about proteins that cannot be identified by peptide mass mapping, or to confirm the identity of proteins that are tentatively identified by an error-tolerant peptide mass search described above. (Griffin et al, Rapid Commun. Mass. Spectrom. 1995, 9, 1546-51).

(b) Analysis of Proteins by Nuclear Magnetic Resonance (NMR)

NMR may be used to characterize the structure of a polypeptide in accordance with the methods of the invention. In particular, NMR can be used, for example, to determine the three dimensional structure, the conformational state, the aggregation level, the state of protein folding/unfolding or the dynamic properties of a polypeptide. For example, the present invention contemplates a method for determining three dimensional structure information of a polypeptide of the invention, the method comprising: (a) generating a purified isotopically labeled polypeptide of the invention; and (b) subjecting the

polypeptide to NMR spectroscopic analysis, thereby determining information about its three dimensional structure.

Interaction between a polypeptide and another molecule can also be monitored using NMR. Thus, the invention encompasses methods for detecting, designing and characterizing interactions between a polypeptide and another molecule, including polypeptides, nucleic acids and small molecules, utilizing NMR techniques. For example, the present invention contemplates a method for determining three dimensional structure information of a polypeptide of the invention, or a fragment thereof, while the polypeptide is complexed with another molecule, the method comprising: (a) generating a purified isotopically labeled polypeptide of the invention, or a fragment thereof; (b) forming a complex between the polypeptide and the other molecule; and (c) subjecting the complex to NMR spectroscopic analysis, thereby determining information about the three dimensional structure of the polypeptide. In another aspect, the present invention contemplates a method for identifying compounds that bind to a polypeptide of the invention, or a fragment thereof, the method comprising: (a) generating a first NMR spectrum of an isotopically labeled polypeptide of the invention, or a fragment thereof; (b) exposing the polypeptide to one or more chemical compounds; (c) generating a second NMR spectrum of the polypeptide which has been exposed to one or more chemical compounds; and (d) comparing the first and second spectra to determine differences between the first and the second spectra, wherein the differences are indicative of one or more compounds that have bound to the polypeptide.

Briefly, the NMR technique involves placing the material to be examined (usually in a suitable solvent) in a powerful magnetic field and irradiating it with radio frequency (rf) electromagnetic radiation. The nuclei of the various atoms will align themselves with the magnetic field until energized by the rf radiation. They then absorb this resonant energy and re-radiate it at a frequency dependent on i) the type of nucleus and ii) its atomic environment. Moreover, resonant energy may be passed from one nucleus to another, either through bonds or through three-dimensional space, thus giving information about the environment of a particular nucleus and nuclei in its vicinity.

However, it is important to recognize that not all nuclei are NMR active. Indeed, not all isotopes of the same element are active. For example, whereas "ordinary" hydrogen, ^1H , is NMR active, heavy hydrogen (deuterium), ^2H , is not active in the same way. Thus, any material that normally contains ^1H hydrogen may be rendered "invisible" in the

hydrogen NMR spectrum by replacing all or almost all the ^1H hydrogens with ^2H . It is for this reason that NMR spectroscopic analyses of water-soluble materials frequently are performed in $^2\text{H}_2\text{O}$ (or deuterium) to eliminate the water signal.

Conversely, "ordinary" carbon, ^{12}C , is NMR inactive whereas the stable isotope, ^{13}C , present to about 1% of total carbon in nature, is active. Similarly, while "ordinary" nitrogen, ^{14}N , is NMR active, it has undesirable properties for NMR and resonates at a different frequency from the stable isotope ^{15}N , present to about 0.4% of total nitrogen in nature.

By labeling proteins with ^{15}N and $^{15}\text{N}/^{13}\text{C}$, it is possible to conduct analytical NMR of macromolecules with weights of 15 kD and 40 kD, respectively. More recently, partial deuteration of the protein in addition to ^{13}C - and ^{15}N -labeling has increased the possible weight of proteins and protein complexes for NMR analysis still further, to approximately 60-70 kD. See Shan et al., *J. Am. Chem. Soc.*, 118:6570-6579 (1996); L.E. Kay, *Methods Enzymol.*, 339:174-203 (2001); and K.H. Gardner & L.E. Kay, *Annu Rev Biophys Biomol Struct.*, 27:357-406 (1998); and references cited therein.

Isotopic substitution may be accomplished by growing a bacterium or yeast or other type of cultured cells, transformed by genetic engineering to produce the protein of choice, in a growth medium containing ^{13}C -, ^{15}N - and/or ^2H -labeled substrates. In certain instances, bacterial growth media consists of ^{13}C -labeled glucose and/or ^{15}N -labeled ammonium salts dissolved in D_2O where necessary. Kay, L. et al., *Science*, 249:411 (1990) and references therein and Bax, A., *J. Am. Chem. Soc.*, 115, 4369 (1993). More recently, isotopically labeled media especially adapted for the labeling of bacterially produced macromolecules have been described. See U.S. Pat. No. 5,324,658.

The goal of these methods has been to achieve universal and/or random isotopic enrichment of all of the amino acids of the protein. By contrast, other methods allow only certain residues to be relatively enriched in ^1H , ^2H , ^{13}C and ^{15}N . For example, Kay et al., *J. Mol. Biol.*, 263, 627-636 (1996) and Kay et al., *J. Am. Chem. Soc.*, 119, 7599-7600 (1997) have described methods whereby isoleucine, alanine, valine and leucine residues in a protein may be labeled with ^2H , ^{13}C and ^{15}N and may be specifically labeled with ^1H at the terminal methyl position. In this way, study of the proton-proton interactions between some amino acids may be facilitated. Similarly, a cell-free system has been described by Yokoyama et al., *J. Biomol. NMR*, 6(2), 129-134 (1995), wherein a transcription-

translation system derived from *E. coli* was used to express human Ha-Ras protein incorporating ^{15}N into serine and/or aspartic acid.

Techniques for producing isotopically labeled proteins and macromolecules, such as glycoproteins, in mammalian or insect cells have been described. See U.S. Pat. Nos. 5,393,669 and 5,627,044; Weller, C. T., *Biochem.*, 35, 8815-23 (1996) and Lustbader, J. W., *J. Biomol. NMR*, 7, 295-304 (1996). Other methods for producing polypeptides and other molecules with labels appropriate for NMR are known in the art.

The present invention contemplates using a variety of solvents which are appropriate for NMR. For ^1H NMR, a deuterium lock solvent may be used. Exemplary deuterium lock solvents include acetone (CD_3COCD_3), chloroform (CDCl_3), dichloro methane (CD_2Cl_2), methylnitrile (CD_3CN), benzene (C_6D_6), water (D_2O), diethylether ($(\text{CD}_3\text{CD}_2)_2\text{O}$), dimethylether ($(\text{CD}_3)_2\text{O}$), *N,N*-dimethylformamide ($(\text{CD}_3)_2\text{NCD O}$), dimethyl sulfoxide (CD_3SOCD_3), ethanol ($\text{CD}_3\text{CD}_2\text{OD}$), methanol (CD_3OD), tetrahydrofuran ($\text{C}_4\text{D}_8\text{O}$), toluene ($\text{C}_6\text{D}_5\text{CD}_3$), pyridine ($\text{C}_5\text{D}_5\text{N}$) and cyclohexane (C_6H_{12}). For example, the present invention contemplates a composition comprising a polypeptide of the invention and a deuterium lock solvent.

The 2-dimensional ^1H - ^{15}N HSQC (Heteronuclear Single Quantum Correlation) spectrum provides a diagnostic fingerprint of conformational state, aggregation level, state of protein folding, and dynamic properties of a polypeptide (Yee et al, *PNAS* 99, 1825-30 (2002)). Polypeptides in aqueous solution usually populate an ensemble of 3-dimensional structures which can be determined by NMR. When the polypeptide is a stable globular protein or domain of a protein, then the ensemble of solution structures is one of very closely related conformations. In this case, one peak is expected for each non-proline residue with a dispersion of resonance frequencies with roughly equal intensity. Additional pairs of peaks from side-chain NH_2 groups are also often observed, and correspond to the approximate number of Gln and Asn residues in the protein. This type of HSQC spectra usually indicates that the protein is amenable to structure determination by NMR methods.

If the HSQC spectrum shows well-dispersed peaks but there are either too few or too many in number, and/or the peak intensities differ throughout the spectrum, then the protein likely does not exist in a single globular conformation. Such spectral features are indicative of conformational heterogeneity with slow or nonexistent inter-conversion between states (too many peaks) or the presence of dynamic processes on an intermediate

timescale that can broaden and obscure the NMR signals. Proteins with this type of spectrum can sometimes be stabilized into a single conformation by changing either the protein construct, the solution conditions, temperature or by binding of another molecule.

The ^1H - ^{15}N HSQC can also indicate whether a protein has formed large nonspecific aggregates or has dynamic properties. Alternatively, proteins that are largely unfolded, e.g., having very little regular secondary structure, result in ^1H - ^{15}N HSQC spectra in which the peaks are all very narrow and intense, but have very little spectral dispersion in the ^{15}N -dimension. This reflects the fact that many or most of the amide groups of amino acids in unfolded polypeptides are solvent exposed and experience similar chemical environments resulting in similar ^1H chemical shifts.

The use of the ^1H - ^{15}N HSQC, can thus allow the rapid characterization of the conformational state, aggregation level, state of protein folding, and dynamic properties of a polypeptide. Additionally, other 2D spectra such as ^1H - ^{13}C HSQC, or HNCQ spectra can also be used in a similar manner. Further use of the ^1H - ^{15}N HSQC combined with relaxation measurements can reveal the molecular rotational correlation time and dynamic properties of polypeptides. The rotational correlation time is proportional to size of the protein and therefore can reveal if it forms specific homo-oligomers such as homodimers, homotetramers, etc.

The structure of stable globular proteins can be determined through a series of well-described procedures. For a general review of structure determination of globular proteins in solution by NMR spectroscopy, see Wüthrich, *Science* 243: 45-50 (1989). See also, Billeter et al., *J. Mol. Biol.* 155: 321-346 (1982). Current methods for structure determination usually require the complete or nearly complete sequence-specific assignment of ^1H -resonance frequencies of the protein and subsequent identification of approximate inter-hydrogen distances (from nuclear Overhauser effect (NOE) spectra) for use in restrained molecular dynamics calculations of the protein conformation. One approach for the analysis of NMR resonance assignments was first outlined by Wüthrich, Wagner and co-workers (Wüthrich, "NMR of proteins and nucleic acids" Wiley, New York, New York (1986); Wüthrich, *Science* 243: 45-50 (1989); Billeter et al., *J. Mol. Biol.* 155: 321-346 (1982)). Newer methods for determining the structures of globular proteins include the use of residual dipolar coupling restraints (Tian et al., *J Am Chem Soc.* 2001 Nov 28;123(47):11791-6; Bax et al, *Methods Enzymol.* 2001;339:127-74) and empirically derived conformational restraints (Zweckstetter & Bax, *J Am Chem Soc.* 2001 Sep

26;123(38):9490-1). It has also been shown that it may be possible to determine structures of globular proteins using only un-assigned NOE measurements. NMR may also be used to determine ensembles of many inter-converting, unfolded conformations (Choy and Forman-Kay, *J Mol Biol.* 2001 May 18;308(5):1011-32).

NMR analysis of a polypeptide in the presence and absence of a test compound (e.g., a polypeptide, nucleic acid or small molecule) may be used to characterize interactions between a polypeptide and another molecule. Because the ^1H - ^{15}N HSQC spectrum and other simple 2D NMR experiments can be obtained very quickly (on the order of minutes depending on protein concentration and NMR instrumentation), they are very useful for rapidly testing whether a polypeptide is able to bind to another molecule. Changes in the resonance frequency (in one or both dimensions) of one or more peaks in the HSQC spectrum indicate an interaction with another molecule. Often only a subset of the peaks will have changes in resonance frequency upon binding to another molecule, allowing one to map onto the structure those residues directly involved in the interaction or involved in conformational changes as a result of the interaction. If the interacting molecule is relatively large (protein or nucleic acid) the peak widths will also broaden due to the increased rotational correlation time of the complex. In some cases the peaks involved in the interaction may actually disappear from the NMR spectrum if the interacting molecule is in intermediate exchange on the NMR timescale (i.e., exchanging on and off the polypeptide at a frequency that is similar to the resonance frequency of the monitored nuclei).

To facilitate the acquisition of NMR data on a large number of compounds (e.g., a library of synthetic or naturally-occurring small organic compounds), a sample changer may be employed. Using the sample changer, a larger number of samples, numbering 60 or more, may be run unattended. To facilitate processing of the NMR data, computer programs are used to transfer and automatically process the multiple one-dimensional NMR data.

In one embodiment, the invention provides a screening method for identifying small molecules capable of interacting with a polypeptide of the invention. In one example, the screening process begins with the generation or acquisition of either a T_2 -filtered or a diffusion-filtered one-dimensional proton spectrum of the compound or mixture of compounds. Means for generating T_2 -filtered or diffusion-filtered one-dimensional proton spectra are well known in the art (see, e.g., S. Meiboom and D. Gill, *Rev. Sci. Instrum.*

29:688(1958), S. J. Gibbs and C. S. Johnson, Jr. J. Main. Reson. 93:395-402 (1991) and A. S. Altieri, et al. J. Am. Chem. Soc. 117: 7566-7567 (1995)).

Following acquisition of the first spectrum for the molecules, the ^{15}N - or ^{13}C -labeled polypeptide is exposed to one or more molecules. Where more than one test compound is to be tested simultaneously, it is preferred to use a library of compounds such as a plurality of small molecules. Such molecules are typically dissolved in perdeuterated dimethylsulfoxide. The compounds in the library may be purchased from vendors or created according to desired needs.

Individual compounds may be selected inter alia on the basis of size and molecular diversity for maximizing the possibility of discovering compounds that interact with widely diverse binding sites of a polypeptide of the invention.

The NMR screening process of the present invention utilizes a range of test compound concentrations, e.g., from about 0.05 to about 1.0 mM. At those exemplary concentrations, compounds which are acidic or basic may significantly change the pH of buffered protein solutions. Chemical shifts are sensitive to pH changes as well as direct binding interactions, and false-positive chemical shift changes, which are not the result of test compound binding but of changes in pH, may therefore be observed. It may therefore be necessary to ensure that the pH of the buffered solution does not change upon addition of the test compound.

Following exposure of the test compounds to a polypeptide (e.g., the target molecule for the experiment) a second one-dimensional T_2 - or diffusion-filtered spectrum is generated. For the T_2 -filtered approach, that second spectrum is generated in the same manner as set forth above. The first and second spectra are then compared to determine whether there are any differences between the two spectra. Differences in the one-dimensional T_2 -filtered spectra indicate that the compound is binding to, or otherwise interacting with, the target molecule. Those differences are determined using standard procedures well known in the art. For the diffusion-filtered method, the second spectrum is generated by looking at the spectral differences between low and high gradient strengths—thus selecting for those compounds whose diffusion rates are comparable to that observed in the absence of target molecule.

To discover additional molecules that bind to the protein, molecules are selected for testing based on the structure/activity relationships from the initial screen and/or structural

information on the initial leads when bound to the protein. By way of example, the initial screening may result in the identification of compounds, all of which contain an aromatic ring. The second round of screening would then use other aromatic molecules as the test compounds.

In another embodiment, the methods of the invention utilize a process for detecting the binding of one ligand to a polypeptide in the presence of a second ligand. In accordance with this embodiment, a polypeptide is bound to the second ligand before exposing the polypeptide to the test compounds.

For more information on NMR methods encompassed by the present invention, see also: U.S. Patent Nos. 5,668,734; 6,194,179; 6,162,627; 6,043,024; 5,817,474; 5,891,642; 5,989,827; 5,891,643; 6,077,682; WO 00/05414; WO 99/22019; Cavanagh, et al., Protein NMR Spectroscopy, Principles and Practice, 1996, Academic Press; Clore, et al., NMR of Proteins. In Topics in Molecular and Structural Biology, 1993, S. Neidle, Fuller, W., and Cohen, J.S., eds., Macmillan Press, Ltd., London; and Christendat et al., Nature Structural Biology 7: 903-909 (2000).

(c) Analysis of Proteins by X-ray Crystallography

(i) X-ray Structure Determination

Exemplary methods for obtaining the three dimensional structure of the crystalline form of a molecule or complex are described herein and, in view of this specification, variations on these methods will be apparent to those skilled in the art (see Ducruix and Geige 1992, IRL Press, Oxford, England).

A variety of methods involving x-ray crystallography are contemplated by the present invention. For example, the present invention contemplates producing a crystallized polypeptide of the invention, or a fragment thereof, by: (a) introducing into a host cell an expression vector comprising a nucleic acid encoding for a polypeptide of the invention, or a fragment thereof; (b) culturing the host cell in a cell culture medium to express the polypeptide or fragment; (c) isolating the polypeptide or fragment from the cell culture; and (d) crystallizing the polypeptide or fragment thereof. Alternatively, the present invention contemplates determining the three dimensional structure of a crystallized polypeptide of the invention, or a fragment thereof, by: (a) crystallizing a polypeptide of the invention, or a fragment thereof, such that the crystals will diffract x-rays to a resolution

of 3.5 Å or better; and (b) analyzing the polypeptide or fragment by x-ray diffraction to determine the three-dimensional structure of the crystallized polypeptide.

X-ray crystallography techniques generally require that the protein molecules be available in the form of a crystal. Crystals may be grown from a solution containing a purified polypeptide of the invention, or a fragment thereof (e.g., a stable domain), by a variety of conventional processes. These processes include, for example, batch, liquid, bridge, dialysis, vapour diffusion (e.g., hanging drop or sitting drop methods). (See for example, McPherson, 1982 John Wiley, New York; McPherson, 1990, *Eur. J. Biochem.* 189: 1-23; Webber, 1991, *Adv. Protein Chem.* 41:1-36).

In certain embodiments, native crystals of the invention may be grown by adding precipitants to the concentrated solution of the polypeptide. The precipitants are added at a concentration just below that necessary to precipitate the protein. Water may be removed by controlled evaporation to produce precipitating conditions, which are maintained until crystal growth ceases.

The formation of crystals is dependent on a number of different parameters, including pH, temperature, protein concentration, the nature of the solvent and precipitant, as well as the presence of added ions or ligands to the protein. In addition, the sequence of the polypeptide being crystallized will have a significant affect on the success of obtaining crystals. Many routine crystallization experiments may be needed to screen all these parameters for the few combinations that might give crystal suitable for x-ray diffraction analysis (See, for example, Jancarik, J & Kim, S.H., *J. Appl. Cryst.* 1991 24: 409-411).

Crystallization robots may automate and speed up the work of reproducibly setting up large number of crystallization experiments. Once some suitable set of conditions for growing the crystal are found, variations of the condition may be systematically screened in order to find the set of conditions which allows the growth of sufficiently large, single, well ordered crystals. In certain instances, a polypeptide of the invention is co-crystallized with a compound that stabilizes the polypeptide.

A number of methods are available to produce suitable radiation for x-ray diffraction. For example, x-ray beams may be produced by synchrotron rings where electrons (or positrons) are accelerated through an electromagnetic field while traveling at close to the speed of light. Because the admitted wavelength may also be controlled, synchrotrons may be used as a tunable x-ray source (Hendrickson WA., *Trends Biochem*

Sci 2000 Dec; 25(12):637-43). For less conventional Laue diffraction studies, polychromatic x-rays covering a broad wavelength window are used to observe many diffraction intensities simultaneously (Stoddard, B. L., Curr. Opin. Struct Biol 1998 Oct; 8(5):612-8). Neutrons may also be used for solving protein crystal structures (Gutberlet T, Heinemann U & Steiner M., Acta Crystallogr D 2001;57: 349-54).

Before data collection commences, a protein crystal may be frozen to protect it from radiation damage. A number of different cryo-protectants may be used to assist in freezing the crystal, such as methyl pentanediol (MPD), isopropanol, ethylene glycol, glycerol, formate, citrate, mineral oil, or a low-molecular-weight polyethylene glycol (PEG). The present invention contemplates a composition comprising a polypeptide of the invention and a cryo-protectant. As an alternative to freezing the crystal, the crystal may also be used for diffraction experiments performed at temperatures above the freezing point of the solution. In these instances, the crystal may be protected from drying out by placing it in a narrow capillary of a suitable material (generally glass or quartz) with some of the crystal growth solution included in order to maintain vapour pressure.

X-ray diffraction results may be recorded by a number of ways known to one of skill in the art. Examples of area electronic detectors include charge coupled device detectors, multi-wire area detectors and phosphorimager detectors (Amemiya, Y, 1997. Methods in Enzymology, Vol. 276. Academic Press, San Diego, pp. 233-243; Westbrook, E. M., Naday, I. 1997. Methods in Enzymology, Vol. 276. Academic Press, San Diego, pp. 244-268; 1997. Kahn, R. & Fourme, R. Methods in Enzymology, Vol. 276. Academic Press, San Diego, pp. 268-286).

A suitable system for laboratory data collection might include a Bruker AXS Proteum R system, equipped with a copper rotating anode source, Confocal Max-Flux™ optics and a SMART 6000 charge coupled device detector. Collection of x-ray diffraction patterns are well documented by those skilled in the art (See, for example, Ducruix and Geige, 1992, IRL Press, Oxford, England).

The theory behind diffraction by a crystal upon exposure to x-rays is well known. Because phase information is not directly measured in the diffraction experiment, and is needed to reconstruct the electron density map, methods that can recover this missing information are required. One method of solving structures *ab initio* are the real / reciprocal space cycling techniques. Suitable real / reciprocal space cycling search

programs include shake-and-bake (Weeks CM, DeTitta GT, Hauptman HA, Thuman P, Miller R *Acta Crystallogr A* 1994; V50: 210-20).

Other methods for deriving phases may also be needed. These techniques generally rely on the idea that if two or more measurements of the same reflection are made where strong, measurable, differences are attributable to the characteristics of a small subset of the atoms alone, then the contributions of other atoms can be, to a first approximation, ignored, and positions of these atoms may be determined from the difference in scattering by one of the above techniques. Knowing the position and scattering characteristics of those atoms, one may calculate what phase the overall scattering must have had to produce the observed differences.

One version of this technique is isomorphous replacement technique, which requires the introduction of new, well ordered, x-ray scatterers into the crystal. These additions are usually heavy metal atoms, (so that they make a significant difference in the diffraction pattern); and if the additions do not change the structure of the molecule or of the crystal cell, the resulting crystals should be isomorphous. Isomorphous replacement experiments are usually performed by diffusing different heavy-metal metals into the channels of a pre-existing protein crystal. Growing the crystal from protein that has been soaked in the heavy atom is also possible (Petsko, G.A., 1985. *Methods in Enzymology*, Vol. 114. Academic Press, Orlando, pp. 147-156). Alternatively, the heavy atom may also be reactive and attached covalently to exposed amino acid side chains (such as the sulfur atom of cysteine) or it may be associated through non-covalent interactions. It is sometimes possible to replace endogenous light metals in metallo-proteins with heavier ones, e.g., zinc by mercury, or calcium by samarium (Petsko, G.A., 1985. *Methods in Enzymology*, Vol. 114. Academic Press, Orlando, pp. 147-156). Exemplary sources for such heavy compounds include, without limitation, sodium bromide, sodium selenate, trimethyl lead acetate, mercuric chloride, methyl mercury acetate, platinum tetracyanide, platinum tetrachloride, nickel chloride, and europium chloride.

A second technique for generating differences in scattering involves the phenomenon of anomalous scattering. X-rays that cause the displacement of an electron in an inner shell to a higher shell are subsequently rescattered, but there is a time lag that shows up as a phase delay. This phase delay is observed as a (generally quite small) difference in intensity between reflections known as Friedel mates that would be identical if no anomalous scattering were present. A second effect related to this phenomenon is that

differences in the intensity of scattering of a given atom will vary in a wavelength dependent manner, given rise to what are known as dispersive differences. In principle anomalous scattering occurs with all atoms, but the effect is strongest in heavy atoms, and may be maximized by using x-rays at a wavelength where the energy is equal to the difference in energy between shells. The technique therefore requires the incorporation of some heavy atom much as is needed for isomorphous replacement, although for anomalous scattering a wider variety of atoms are suitable, including lighter metal atoms (copper, zinc, iron) in metallo-proteins. One method for preparing a protein for anomalous scattering involves replacing the methionine residues in whole or in part with selenium containing seleno-methionine. Soaks with halide salts such as bromides and other non-reactive ions may also be effective (Dauter Z, Li M, Wlodawer A., *Acta Crystallogr D* 2001; 57: 239-49).

In another process, known as multiple anomalous scattering or MAD, two to four suitable wavelengths of data are collected. (Hendrickson, W.A. and Ogata, C.M. 1997 *Methods in Enzymology* 276, 494 – 523). Phasing by various combinations of single and multiple isomorphous and anomalous scattering are possible too. For example, SIRAS (single isomorphous replacement with anomalous scattering) utilizes both the isomorphous and anomalous differences for one derivative to derive phases. More traditionally, several different heavy atoms are soaked into different crystals to get sufficient phase information from isomorphous differences while ignoring anomalous scattering, in the technique known as multiple isomorphous replacement (MIR) (Petsko, G.A., 1985. *Methods in Enzymology*, Vol. 114. Academic Press, Orlando, pp. 147-156).

Additional restraints on the phases may be derived from density modification techniques. These techniques use either generally known features of electron density distribution or known facts about that particular crystal to improve the phases. For example, because protein regions of the crystal scatter more strongly than solvent regions, solvent flattening/flipping may be used to adjust phases to make solvent density a uniform flat value (Zhang, K. Y. J., Cowtan, K. and Main, P. *Methods in Enzymology* 277, 1997 Academic Press, Orlando pp 53-64). If more than one molecule of the protein is present in the asymmetric unit, the fact that the different molecules should be virtually identical may be exploited to further reduce phase error using non-crystallographic symmetry averaging (Villieux, F. M. D. and Read, R. J. *Methods in Enzymology* 277, 1997 Academic Press, Orlando pp18-52). Suitable programs for performing these processes include DM and other

programs of the CCP4 suite (Collaborative Computational Project, Number 4, 1994, *Acta Cryst. D* 50, 760-763) and CNX.

The unit cell dimensions, symmetry, vector amplitude and derived phase information can be used in a Fourier transform function to calculate the electron density in the unit cell, i.e., to generate an experimental electron density map. This may be accomplished using programs of the CNX or CCP4 packages. The resolution is measured in Angstrom (\AA) units, and is closely related to how far apart two objects need to be before they can be reliably distinguished. The smaller this number is, the higher the resolution and therefore the greater the amount of detail that can be seen. Preferably, crystals of the invention diffract x-rays to a resolution of better than about 4.0, 3.5, 3.0, 2.5, 2.0, 1.5, 1.0, 0.5 \AA or better.

As used herein, the term "modeling" includes the quantitative and qualitative analysis of molecular structure and/or function based on atomic structural information and interaction models. The term "modeling" includes conventional numeric-based molecular dynamic and energy minimization models, interactive computer graphic models, modified molecular mechanics models, distance geometry and other structure-based constraint models.

Model building may be accomplished by either the crystallographer using a computer graphics program such as TURBO or O (Jones, T.A. et al., *Acta Crystallogr. A* 47, 100-119, 1991) or, under suitable circumstances, by using a fully automated model building program, such as wARP (Anastassis Perrakis, Richard Morris & Victor S. Lamzin; *Nature Structural Biology*, May 1999 Volume 6 Number 5 pp 458 - 463) or MAID (Levitt, D. G., *Acta Crystallogr. D* 2001 V57: 1013-9). This structure may be used to calculate model-derived diffraction amplitudes and phases. The model-derived and experimental diffraction amplitudes may be compared and the agreement between them can be described by a parameter referred to as R-factor. A high degree of correlation in the amplitudes corresponds to a low R-factor value, with 0.0 representing exact agreement and 0.59 representing a completely random structure. Because the R-factor may be lowered by introducing more free parameters into the model, an unbiased, cross-correlated version of the R-factor known as the R-free gives a more objective measure of model quality. For the calculation of this parameter a subset of reflections (generally around 10%) are set aside at the beginning of the refinement and not used as part of the refinement target. These

reflections are then compared to those predicted by the model (Kleywegt GJ, Brunger AT, Structure 1996 Aug 15;4(8):897-904).

The model may be improved using computer programs that maximize the probability that the observed data was produced from the predicted model, while simultaneously optimizing the model geometry. For example, the CNX program may be used for model refinement, as can the XPLOR program (1992, Nature 355:472-475, G.N. Murshudov, A.A.Vagin and E.J.Dodson, (1997) Acta Cryst. D 53, 240-255). In order to maximize the convergence radius of refinement, simulated annealing refinement using torsion angle dynamics may be employed in order to reduce the degrees of freedom of motion of the model (Adams PD, Pannu NS, Read RJ, Brunger AT., Proc Natl Acad Sci U S A 1997 May 13;94(10):5018-23). Where experimental phase information is available (e.g. where MAD data was collected) Hendrickson-Lattman phase probability targets may be employed. Isotropic or anisotropic domain, group or individual temperature factor refinement, may be used to model variance of the atomic position from its mean. Well defined peaks of electron density not attributable to protein atoms are generally modeled as water molecules. Water molecules may be found by manual inspection of electron density maps, or with automatic water picking routines. Additional small molecules, including ions, cofactors, buffer molecules or substrates may be included in the model if sufficiently unambiguous electron density is observed in a map.

In general, the R-free is rarely as low as 0.15 and may be as high as 0.35 or greater for a reasonably well-determined protein structure. The residual difference is a consequence of approximations in the model (inadequate modeling of residual structure in the solvent, modeling atoms as isotropic Gaussian spheres, assuming all molecules are identical rather than having a set of discrete conformers, etc.) and errors in the data (Lattman EE., Proteins 1996; 25: i-ii). In refined structures at high resolution, there are usually no major errors in the orientation of individual residues, and the estimated errors in atomic positions are usually around 0.1 - 0.2 up to 0.3 Å.

The three dimensional structure of a new crystal may be modeled using molecular replacement. The term "molecular replacement" refers to a method that involves generating a preliminary model of a molecule or complex whose structure coordinates are unknown, by orienting and positioning a molecule whose structure coordinates are known within the unit cell of the unknown crystal, so as best to account for the observed diffraction pattern of the unknown crystal. Phases may then be calculated from this model and combined with

the observed amplitudes to give an approximate Fourier synthesis of the structure whose coordinates are unknown. This, in turn, can be subject to any of the several forms of refinement to provide a final, accurate structure of the unknown crystal. Lattman, E., "Use of the Rotation and Translation Functions", in *Methods in Enzymology*, 115, pp. 55-77 (1985); M. G. Rossmann, ed., "The Molecular Replacement Method", *Int. Sci. Rev. Ser.*, No. 13, Gordon & Breach, New York, (1972).

Commonly used computer software packages for molecular replacement are CNX, X-PLOR (Brunger 1992, *Nature* 355: 472-475), AMoRE (Navaza, 1994, *Acta Crystallogr. A* 50:157-163), the CCP4 package, the MERLOT package (P.M.D. Fitzgerald, *J. Appl. Cryst.*, Vol. 21, pp. 273-278, 1988) and XTALVIEW (McCree et al (1992) *J. Mol. Graphics* 10: 44-46). The quality of the model may be analyzed using a program such as PROCHECK or 3D-Profler (Laskowski et al 1993 *J. Appl. Cryst.* 26:283-291; Luthy R. et al, *Nature* 356: 83-85, 1992; and Bowie, J.U. et al, *Science* 253: 164-170, 1991).

Homology modeling (also known as comparative modeling or knowledge-based modeling) methods may also be used to develop a three dimensional model from a polypeptide sequence based on the structures of known proteins. The method utilizes a computer model of a known protein, a computer representation of the amino acid sequence of the polypeptide with an unknown structure, and standard computer representations of the structures of amino acids. This method is well known to those skilled in the art (Greer, 1985, *Science* 228, 1055; Bundell et al 1988, *Eur. J. Biochem.* 172, 513; Knighton et al., 1992, *Science* 258:130-135, <http://biochem.vt.edu/courses/-modeling/homology.htm>). Computer programs that can be used in homology modeling are QUANTA and the Homology module in the Insight II modeling package distributed by Molecular Simulations Inc, or MODELLER (Rockefeller University, www.iucr.ac.uk/sinris-top/logical/prg-modeller.html).

Once a homology model has been generated it is analyzed to determine its correctness. A computer program available to assist in this analysis is the Protein Health module in QUANTA which provides a variety of tests. Other programs that provide structure analysis along with output include PROCHECK and 3D-Profler (Luthy R. et al, *Nature* 356: 83-85, 1992; and Bowie, J.U. et al, *Science* 253: 164-170, 1991). Once any irregularities have been resolved, the entire structure may be further refined.

Other molecular modeling techniques may also be employed in accordance with this invention. See, e.g., Cohen, N. C. *et al*, J. Med. Chem., 33, pp. 883-894 (1990). See also, Navix, M. A. and M. A. Marko, Current Opinions in Structural Biology, 2, pp. 202-210 (1992).

Under suitable circumstances, the entire process of solving a crystal structure may be accomplished in an automated fashion by a system such as ELVES (<http://ucxray.berkeley.edu/~jamesh/elves/index.html>) with little or no user intervention.

(ii) X-ray Structure

The present invention provides methods for determining some or all of the structural coordinates for amino acids of a polypeptide of the invention, or a complex thereof.

In another aspect, the present invention provides methods for identifying a druggable region of a polypeptide of the invention. For example, one such method includes: (a) obtaining crystals of a polypeptide of the invention or a fragment thereof such that the three dimensional structure of the crystallized protein can be determined to a resolution of 3.5 Å or better; (b) determining the three dimensional structure of the crystallized polypeptide or fragment using x-ray diffraction; and (c) identifying a druggable region of a polypeptide of the invention based on the three-dimensional structure of the polypeptide or fragment.

A three dimensional structure of a molecule or complex may be described by the set of atoms that best predict the observed diffraction data (that is, which possesses a minimal R value). Files may be created for the structure that defines each atom by its chemical identity, spatial coordinates in three dimensions, root mean squared deviation from the mean observed position and fractional occupancy of the observed position.

Those of skill in the art understand that a set of structure coordinates for an protein, complex or a portion thereof, is a relative set of points that define a shape in three dimensions. Thus, it is possible that an entirely different set of coordinates could define a similar or identical shape. Moreover, slight variations in the individual coordinates may have little affect on overall shape. Such variations in coordinates may be generated because of mathematical manipulations of the structure coordinates. For example, structure coordinates could be manipulated by crystallographic permutations of the structure coordinates, fractionalization of the structure coordinates, integer additions or subtractions to sets of the structure coordinates, inversion of the structure coordinates or any

combination of the above. Alternatively, modifications in the crystal structure due to mutations, additions, substitutions, and/or deletions of amino acids, or other changes in any of the components that make up the crystal, could also yield variations in structure coordinates. Such slight variations in the individual coordinates will have little effect on overall shape. If such variations are within an acceptable standard error as compared to the original coordinates, the resulting three-dimensional shape is considered to be structurally equivalent. It should be noted that slight variations in individual structure coordinates of a polypeptide of the invention or a complex thereof would not be expected to significantly alter the nature of modulators that could associate with a druggable region thereof. Thus, for example, a modulator that bound to the active site of a polypeptide of the invention would also be expected to bind to or interfere with another active site whose structure coordinates define a shape that falls within the acceptable error.

A crystal structure of the present invention may be used to make a structural or computer model of the polypeptide, complex or portion thereof. A model may represent the secondary, tertiary and/or quaternary structure of the polypeptide, complex or portion. The configurations of points in space derived from structure coordinates according to the invention can be visualized as, for example, a holographic image, a stereodiagram, a model or a computer-displayed image, and the invention thus includes such images, diagrams or models.

(iii) Structural Equivalents

Various computational analyses can be used to determine whether a molecule or the active site portion thereof is structurally equivalent with respect to its three-dimensional structure, to all or part of a structure of a polypeptide of the invention or a portion thereof.

For the purpose of this invention, any molecule or complex or portion thereof, that has a root mean square deviation of conserved residue backbone atoms (N, C α , C, O) of less than about 1.75 Å, when superimposed on the relevant backbone atoms described by the reference structure coordinates of a polypeptide of the invention, is considered "structurally equivalent" to the reference molecule. That is to say, the crystal structures of those portions of the two molecules are substantially identical, within acceptable error. Alternatively, the root mean square deviation may be less than about 1.50, 1.40, 1.25, 1.0, 0.75, 0.5 or 0.35 Å.

The term "root mean square deviation" is understood in the art and means the square root of the arithmetic mean of the squares of the deviations. It is a way to express the deviation or variation from a trend or object.

In another aspect, the present invention provides a scalable three-dimensional configuration of points, at least a portion of said points, and preferably all of said points, derived from structural coordinates of at least a portion of a polypeptide of the invention and having a root mean square deviation from the structure coordinates of the polypeptide of the invention of less than 1.50, 1.40, 1.25, 1.0, 0.75, 0.5 or 0.35 Å. In certain embodiments, the portion of a polypeptide of the invention is 25%, 33%, 50%, 66%, 75%, 85%, 90% or 95% or more of the amino acid residues contained in the polypeptide.

In another aspect, the present invention provides a molecule or complex including a druggable region of a polypeptide of the invention, the druggable region being defined by a set of points having a root mean square deviation of less than about 1.75 Å from the structural coordinates for points representing (a) the backbone atoms of the amino acids contained in a druggable region of a polypeptide of the invention, (b) the side chain atoms (and optionally the C α atoms) of the amino acids contained in such druggable region, or (c) all the atoms of the amino acids contained in such druggable region. In certain embodiments, only a portion of the amino acids of a druggable region may be included in the set of points, such as 25%, 33%, 50%, 66%, 75%, 85%, 90% or 95% or more of the amino acid residues contained in the druggable region. In certain embodiments, the root mean square deviation may be less than 1.50, 1.40, 1.25, 1.0, 0.75, 0.5, or 0.35 Å. In still other embodiments, instead of a druggable region, a stable domain, fragment or structural motif is used in place of a druggable region.

(iv) Machine Displays and Machine Readable Storage Media

The invention provides a machine-readable storage medium including a data storage material encoded with machine readable data which, when using a machine programmed with instructions for using said data, displays a graphical three-dimensional representation of any of the molecules or complexes, or portions thereof, of this invention. In another embodiment, the graphical three-dimensional representation of such molecule, complex or portion thereof includes the \pm a root mean square deviation of certain atoms of such molecule by a specified amount, such as the backbone atoms by less than 0.8 Å. In another embodiment, a structural equivalent of such molecule, complex, or portion thereof, may be

displayed. In another embodiment, the portion may include a druggable region of the polypeptide of the invention.

According to one embodiment, the invention provides a computer for determining at least a portion of the structure coordinates corresponding to x-ray diffraction data obtained from a molecule or complex, wherein said computer includes: (a) a machine-readable data storage medium comprising a data storage material encoded with machine-readable data, wherein said data comprises at least a portion of the structural coordinates of a polypeptide of the invention; (b) a machine-readable data storage medium comprising a data storage material encoded with machine-readable data, wherein said data comprises x-ray diffraction data from said molecule or complex; (c) a working memory for storing instructions for processing said machine-readable data of (a) and (b); (d) a central-processing unit coupled to said working memory and to said machine-readable data storage medium of (a) and (b) for performing a Fourier transform of the machine readable data of (a) and for processing said machine readable data of (b) into structure coordinates; and (e) a display coupled to said central-processing unit for displaying said structure coordinates of said molecule or complex. In certain embodiments, the structural coordinates displayed are structurally equivalent to the structural coordinates of a polypeptide of the invention.

In an alternative embodiment, the machine-readable data storage medium includes a data storage material encoded with a first set of machine readable data which includes the Fourier transform of the structure coordinates of a polypeptide of the invention or a portion thereof, and which, when using a machine programmed with instructions for using said data, can be combined with a second set of machine readable data including the x-ray diffraction pattern of a molecule or complex to determine at least a portion of the structure coordinates corresponding to the second set of machine readable data.

For example, a system for reading a data storage medium may include a computer including a central processing unit ("CPU"), a working memory which may be, e.g., RAM (random access memory) or "core" memory, mass storage memory (such as one or more disk drives or CD-ROM drives), one or more display devices (e.g., cathode-ray tube ("CRT") displays, light emitting diode ("LED") displays, liquid crystal displays ("LCDs"), electroluminescent displays, vacuum fluorescent displays, field emission displays ("FEDs"), plasma displays, projection panels, etc.), one or more user input devices (e.g., keyboards, microphones, mice, touch screens, etc.), one or more input lines, and one or more output lines, all of which are interconnected by a conventional bidirectional system

bus. The system may be a stand-alone computer, or may be networked (e.g., through local area networks, wide area networks, intranets, extranets, or the internet) to other systems (e.g., computers, hosts, servers, etc.). The system may also include additional computer controlled devices such as consumer electronics and appliances.

Input hardware may be coupled to the computer by input lines and may be implemented in a variety of ways. Machine-readable data of this invention may be inputted via the use of a modem or modems connected by a telephone line or dedicated data line. Alternatively or additionally, the input hardware may include CD-ROM drives or disk drives. In conjunction with a display terminal, a keyboard may also be used as an input device.

Output hardware may be coupled to the computer by output lines and may similarly be implemented by conventional devices. By way of example, the output hardware may include a display device for displaying a graphical representation of an active site of this invention using a program such as QUANTA as described herein. Output hardware might also include a printer, so that hard copy output may be produced, or a disk drive, to store system output for later use.

In operation, a CPU coordinates the use of the various input and output devices, coordinates data accesses from mass storage devices, accesses to and from working memory, and determines the sequence of data processing steps. A number of programs may be used to process the machine-readable data of this invention. Such programs are discussed in reference to the computational methods of drug discovery as described herein. References to components of the hardware system are included as appropriate throughout the following description of the data storage medium.

Machine-readable storage devices useful in the present invention include, but are not limited to, magnetic devices, electrical devices, optical devices, and combinations thereof. Examples of such data storage devices include, but are not limited to, hard disk devices, CD devices, digital video disk devices, floppy disk devices, removable hard disk devices, magneto-optic disk devices, magnetic tape devices, flash memory devices, bubble memory devices, holographic storage devices, and any other mass storage peripheral device. It should be understood that these storage devices include necessary hardware (e.g., drives, controllers, power supplies, etc.) as well as any necessary media (e.g., disks, flash cards, etc.) to enable the storage of data..

In one embodiment, the present invention contemplates a computer readable storage medium comprising structural data, wherein the data include the identity and three-dimensional coordinates of a polypeptide of the invention or portion thereof. In another aspect, the present invention contemplates a database comprising the identity and three-dimensional coordinates of a polypeptide of the invention or a portion thereof. Alternatively, the present invention contemplates a database comprising a portion or all of the atomic coordinates of a polypeptide of the invention or portion thereof.

(v) Structurally Similar Molecules and Complexes

Structural coordinates for a polypeptide of the invention can be used to aid in obtaining structural information about another molecule or complex. This method of the invention allows determination of at least a portion of the three-dimensional structure of molecules or molecular complexes which contain one or more structural features that are similar to structural features of a polypeptide of the invention. Similar structural features can include, for example, regions of amino acid identity, conserved active site or binding site motifs, and similarly arranged secondary structural elements (e.g., α helices and β sheets). Many of the methods described above for determining the structure of a polypeptide of the invention may be used for this purpose as well.

For the present invention, a "structural homolog" is a polypeptide that contains one or more amino acid substitutions, deletions, additions, or rearrangements with respect to the amino acid sequence of SEQ ID NO: 4 or other polypeptide of the invention, but that, when folded into its native conformation, exhibits or is reasonably expected to exhibit at least a portion of the tertiary (three-dimensional) structure of the polypeptide encoded by SEQ ID NO: 4 or such other polypeptide of the invention. For example, structurally homologous molecules can contain deletions or additions of one or more contiguous or noncontiguous amino acids, such as a loop or a domain. Structurally homologous molecules also include modified polypeptide molecules that have been chemically or enzymatically derivatized at one or more constituent amino acids, including side chain modifications, backbone modifications, and N- and C-terminal modifications including acetylation, hydroxylation, methylation, amidation, and the attachment of carbohydrate or lipid moieties, cofactors, and the like.

By using molecular replacement, all or part of the structure coordinates of a polypeptide of the invention can be used to determine the structure of a crystallized

molecule or complex whose structure is unknown more quickly and efficiently than attempting to determine such information *ab initio*. For example, in one embodiment this invention provides a method of utilizing molecular replacement to obtain structural information about a molecule or complex whose structure is unknown including: (a) crystallizing the molecule or complex of unknown structure; (b) generating an x-ray diffraction pattern from said crystallized molecule or complex; and (c) applying at least a portion of the structure coordinates for a polypeptide of the invention to the x-ray diffraction pattern to generate a three-dimensional electron density map of the molecule or complex whose structure is unknown.

In another aspect, the present invention provides a method for generating a preliminary model of a molecule or complex whose structure coordinates are unknown, by orienting and positioning the relevant portion of a polypeptide of the invention within the unit cell of the crystal of the unknown molecule or complex so as best to account for the observed x-ray diffraction pattern of the crystal of the molecule or complex whose structure is unknown.

Structural information about a portion of any crystallized molecule or complex that is sufficiently structurally similar to a portion of a polypeptide of the invention may be resolved by this method. In addition to a molecule that shares one or more structural features with a polypeptide of the invention, a molecule that has similar bioactivity, such as the same catalytic activity, substrate specificity or ligand binding activity as a polypeptide of the invention, may also be sufficiently structurally similar to a polypeptide of the invention to permit use of the structure coordinates for a polypeptide of the invention to solve its crystal structure.

In another aspect, the method of molecular replacement is utilized to obtain structural information about complex containing a polypeptide of the invention, such as a complex between a modulator and a polypeptide of the invention (or a domain, fragment, ortholog, homolog etc. thereof). In certain instances, the complex includes a polypeptide of the invention (or a domain, fragment, ortholog, homolog etc. thereof) co-complexed with a modulator. For example, in one embodiment, the present invention contemplates a method for making a crystallized complex comprising a polypeptide of the invention, or a fragment thereof, and a compound having a molecular weight of less than 5 kDa, the method comprising: (a) crystallizing a polypeptide of the invention such that the crystals will diffract x-rays to a resolution of 3.5 Å or better; and (b) soaking the crystal in a solution

comprising the compound having a molecular weight of less than 5 kDa, thereby producing a crystallized complex comprising the polypeptide and the compound.

Using homology modeling, a computer model of a structural homolog or other polypeptide can be built or refined without crystallizing the molecule. For example, in another aspect, the present invention provides a computer-assisted method for homology modeling a structural homolog of a polypeptide of the invention including: aligning the amino acid sequence of a known or suspected structural homolog with the amino acid sequence of a polypeptide of the invention and incorporating the sequence of the homolog into a model of a polypeptide of the invention derived from atomic structure coordinates to yield a preliminary model of the homolog; subjecting the preliminary model to energy minimization to yield an energy minimized model; remodeling regions of the energy minimized model where stereochemistry restraints are violated to yield a final model of the homolog.

In another embodiment, the present invention contemplates a method for determining the crystal structure of a homolog of a polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, or equivalent thereof, the method comprising: (a) providing the three dimensional structure of a crystallized polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, or a fragment thereof; (b) obtaining crystals of a homologous polypeptide comprising an amino acid sequence that is at least 80% identical to the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4 such that the three dimensional structure of the crystallized homologous polypeptide may be determined to a resolution of 3.5 Å or better; and (c) determining the three dimensional structure of the crystallized homologous polypeptide by x-ray crystallography based on the atomic coordinates of the three dimensional structure provided in step (a). In certain instances of the foregoing method, the atomic coordinates for the homologous polypeptide have a root mean square deviation from the backbone atoms of the polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, or a fragment thereof, of not more than 1.5 Å for all backbone atoms shared in common with the homologous polypeptide and the polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, or a fragment thereof.

(vi) NMR Analysis Using X-ray Structural Data

In another aspect, the structural coordinates of a known crystal structure may be applied to nuclear magnetic resonance data to determine the three dimensional structures of

polypeptides with uncharacterized or incompletely characterized structure. (See for example, Wuthrich, 1986, John Wiley and Sons, New York: 176-199; Pflugrath et al., 1986, J. Molecular Biology 189: 383-386; Kline et al., 1986 J. Molecular Biology 189:377-382). While the secondary structure of a polypeptide may often be determined by NMR data, the spatial connections between individual pieces of secondary structure are not as readily determined. The structural coordinates of a polypeptide defined by x-ray crystallography can guide the NMR spectroscopist to an understanding of the spatial interactions between secondary structural elements in a polypeptide of related structure. Information on spatial interactions between secondary structural elements can greatly simplify NOE data from two-dimensional NMR experiments. In addition, applying the structural coordinates after the determination of secondary structure by NMR techniques simplifies the assignment of NOE's relating to particular amino acids in the polypeptide sequence.

In an embodiment, the invention relates to a method of determining three dimensional structures of polypeptides with unknown structures, by applying the structural coordinates of a crystal of the present invention to nuclear magnetic resonance data of the unknown structure. This method comprises the steps of: (a) determining the secondary structure of an unknown structure using NMR data; and (b) simplifying the assignment of through-space interactions of amino acids. The term "through-space interactions" defines the orientation of the secondary structural elements in the three dimensional structure and the distances between amino acids from different portions of the amino acid sequence. The term "assignment" defines a method of analyzing NMR data and identifying which amino acids give rise to signals in the NMR spectrum.

For all of this section on x-ray crystallography, see also Brooks et al. (1983) *J Comput Chem* 4:187-217; Weiner et al (1981) *J. Comput. Chem.* 106: 765; Eisenfield et al. (1991) *Am J Physiol* 261:C376-386; Lybrand (1991) *J Pharm Belg* 46:49-54; Froimowitz (1990) *Biotechniques* 8:640-644; Burbam et al. (1990) *Proteins* 7:99-111; Pedersen (1985) *Environ Health Perspect* 61:185-190; and Kini et al. (1991) *J Biomol Struct Dyn* 9:475-488; Ryckaert et al. (1977) *J Comput Phys* 23:327; Van Gunsteren et al. (1977) *Mol Phys* 34:1311; Anderson (1983) *J Comput Phys* 52:24; J. Mol. Biol. 48: 442-453, 1970; Dayhoff et al., *Meth. Enzymol.* 91: 524-545, 1983; Henikoff and Henikoff, *Proc. Nat. Acad. Sci. USA* 89: 10915-10919, 1992; J. Mol. Biol. 233: 716-738, 1993; *Methods in Enzymology*, Volume 276, *Macromolecular crystallography, Part A*, ISBN 0-12-182177-3 and Volume 277, *Macromolecular crystallography, Part B*, ISBN 0-12-182178-1, Eds. Charles W.

Carter, Jr. and Robert M. Sweet (1997), Academic Press, San Diego; Pfuetzner, et al., J. Biol. Chem. 272: 430-434 (1997).

6. Interacting Proteins

The present invention also provides methods for isolating specific protein interactors of a polypeptide of the invention, and complexes comprising a polypeptide of the invention and one or more interacting proteins. In one aspect, the present invention contemplates an isolated protein complex comprising a polypeptide of the invention and at least one protein that interacts with the polypeptide of the invention. The protein may be naturally-occurring. The interacting protein may be of *P. aeruginosa* origin. Alternatively, the interacting protein may be of mammalian origin or human origin. Either the polypeptide of the invention or the interacting protein or both may be a fusion protein.

The present invention contemplates a method for identifying a protein capable of interacting with a polypeptide of the invention or a fragment thereof, the method comprising: (a) exposing a sample to a solid substrate coupled to a polypeptide of the invention or a fragment thereof under conditions which promote protein-protein interactions; (b) washing the solid substrate so as to remove any polypeptides interacting non-specifically with the polypeptide or fragment; (c) eluting the polypeptides which specifically interact with the polypeptide or fragment; and (d) identifying the interacting protein. The sample may be an extract of *P. aeruginosa*, a mammalian cell extract, a human cell extract, a purified protein (or a fragment thereof), or a mixture of purified proteins (or fragments thereof). The interacting protein may be identified by a number of methods, including mass spectrometry or protein sequencing.

In another aspect, the present invention contemplates a method for identifying a protein capable of interacting with a polypeptide of present invention or a fragment thereof, the method comprising: (a) subjecting a sample to protein-affinity chromatography on multiple columns, the columns having a polypeptide of the invention or a fragment thereof coupled to the column matrix in varying concentrations, and eluting bound components of the extract from the columns; (b) separating the components to isolate a polypeptide capable of interacting with the polypeptide or fragment; and (c) analyzing the interacting protein by mass spectrometry to identify the interacting protein. In certain instances, the foregoing method will use polyacrylamide gel electrophoresis without SDS.

In another aspect, the present invention contemplates a method for identifying a protein capable of interacting with a polypeptide of the invention, the method comprising: (a) subjecting a cellular extract or extracellular fluid to protein-affinity chromatography on multiple columns, the columns having a polypeptide of the invention or a fragment thereof coupled to the column matrix in varying concentrations, and eluting bound components of the extract from the columns; (b) gel-separating the components to isolate an interacting protein; wherein the interacting protein is observed to vary in amount in direct relation to the concentration of coupled polypeptide or fragment; (c) digesting the interacting protein to give corresponding peptides; (d) analyzing the peptides by MALDI-TOF mass spectrometry or post source decay to determine the peptide masses; and (e) performing correlative database searches with the peptide, or peptide fragment, masses, whereby the interacting protein is identified based on the masses of the peptides or peptide fragments. The foregoing method may include the further step of including the identifies of any interacting proteins into a relational database.

In another aspect, the invention further contemplates a method for identifying modulators of a protein complex, the method comprising: (a) contacting a protein complex comprising a polypeptide of the invention and an interacting protein with one or more test compounds; and (b) determining the effect of the test compound on (i) the activity of the protein complex, (ii) the amount of the protein complex, (iii) the stability of the protein complex, (iv) the conformation of the protein complex, (v) the activity of at least one polypeptide included in the protein complex, (vi) the conformation of at least one polypeptide included in the protein complex, (vii) the intracellular localization of the protein complex or a component thereof, (viii) the transcription level of a gene dependent on the complex, and/or (ix) the level of second messenger levels in a cell; thereby identifying modulators of the protein complex. The foregoing method may be carried out *in vitro* or *in vivo* as appropriate.

Typically, it will be desirable to immobilize a polypeptide of the invention to facilitate separation of complexes comprising a polypeptide of the invention from uncomplexed forms of the interacting proteins, as well as to accommodate automation of the assay. The polypeptide of the invention, or ligand, may be immobilized onto a solid support (e.g., column matrix, microtiter plate, slide, etc.). In certain embodiments, the ligand may be purified. In certain instances, a fusion protein may be provided which adds a domain that permits the ligand to be bound to a support.

In various *in vitro* embodiments, the set of proteins engaged in a protein-protein interaction comprises a cell extract, a clarified cell extract, or a reconstituted protein mixture of at least semi-purified proteins. By semi-purified, it is meant that the proteins utilized in the reconstituted mixture have been previously separated from other cellular or viral proteins. For instance, in contrast to cell lysates, the proteins involved in a protein-protein interaction are present in the mixture to at least about 50% purity relative to all other proteins in the mixture, and more preferably are present in greater, even 90-95%, purity. In certain embodiments of the subject method, the reconstituted protein mixture is derived by mixing highly purified proteins such that the reconstituted mixture substantially lacks other proteins (such as of cellular or viral origin) which might interfere with or otherwise alter the ability to measure activity resulting from the given protein-protein interaction.

Complex formation involving a polypeptide of the invention and another component polypeptide or a substrate polypeptide, may be detected by a variety of techniques. For instance, modulation in the formation of complexes can be quantitated using, for example, detectably labeled proteins (e.g. radiolabeled, fluorescently labeled, or enzymatically labeled), by immunoassay, or by chromatographic detection.

The present invention also provides assays for identifying molecules which are modulators of a protein-protein interaction involving a polypeptide of the invention, or are a modulator of the role of the complex comprising a polypeptide of the invention in the infectivity or pathogenicity of *P. aeruginosa*. In one embodiment, the assay detects agents which inhibit formation or stabilization of a protein complex comprising a polypeptide of the invention and one or more additional proteins. In another embodiment, the assay detects agents which modulate the intrinsic biological activity of a protein complex comprising a polypeptide of the invention, such as an enzymatic activity, binding to other cellular components, cellular compartmentalization, signal transduction, and the like. Such modulators may be used, for example, in the treatment of *P. aeruginosa* related diseases or disorders. In certain embodiments, the compound is a mechanism based inhibitor which chemically alters one member of a protein-protein interaction involving a polypeptide of the invention and which is a specific inhibitor of that member, e.g. has an inhibition constant about 10-fold, 100-fold, or 1000-fold different compared to homologous proteins.

In one embodiment, proteins that interact with a polypeptide of the invention may be isolated using immunoprecipitation. A polypeptide of the invention may be expressed in

P. aeruginosa, or in a heterologous system. The cells expressing a polypeptide of the invention are then lysed under conditions which maintain protein-protein interactions, and complexes comprising a polypeptide of the invention are isolated. For example, a polypeptide of the invention may be expressed in mammalian cells, including human cells, in order to identify mammalian proteins that interact with a polypeptide of the invention and therefore may play a role in *P. aeruginosa* infectivity or proliferation. In one embodiment, a polypeptide of the invention is expressed in the cell type for which it is desirable to find interacting proteins. For example, a polypeptide of the invention may be expressed in *P. aeruginosa* in order to find *P. aeruginosa* derived interacting proteins.

In an alternative embodiment, a polypeptide of the invention is expressed and purified and then mixed with a potential interacting protein or mixture of proteins to identify complex formation. The potential interacting protein may be a single purified or semi-purified protein, or a mixture of proteins, including a mixture of purified or semi-purified proteins, a cell lysate, a clarified cell lysate, a semi-purified cell lysate, etc.

In certain embodiments, it may be desirable to use a tagged version of a polypeptide of the invention in order to facilitate isolation of complexes from the reaction mixture. Suitable tags for immunoprecipitation experiments include HA, myc, FLAG, HIS, GST, protein A, protein G, etc. Immunoprecipitation from a cell lysate or other protein mixture may be carried out using an antibody specific for a polypeptide of the invention or using an antibody which recognizes a tag to which a polypeptide of the invention is fused (e.g., anti-HA, anti-myc, anti-FLAG, etc.). Antibodies specific for a variety of tags are known to the skilled artisan and are commercially available from a number of sources. In the case where a polypeptide of the invention is fused to a His, GST, or protein A/G tag, immunoprecipitation may be carried out using the appropriate affinity resin (e.g., beads functionalized with Ni, glutathione, Fc region of IgG, etc.). Test compounds which modulate a protein-protein interaction involving a polypeptide of the invention may be identified by carrying out the immunoprecipitation reaction in the presence and absence of the test agent and comparing the level and/or activity of the protein complex between the two reactions.

In another embodiment, proteins that interact with a polypeptide of the invention may be identified using affinity chromatography. Some examples of such chromatography are described in USSN 09/727,812, filed November 30, 2000, and the PCT Application

filed November 30, 2001 and entitled "Methods for Systematic Identification of Protein-Protein Interactions and other Properties", which claims priority to such U.S. application.

In one aspect, for affinity chromatography using a solid support, a polypeptide of the invention or a fragment thereof may be attached by a variety of means known to those of skill in the art. For example, the polypeptide may be coupled directly (through a covalent linkage) to commercially available pre-activated resins as described in Formosa et al., *Methods in Enzymology* 1991, 208, 24-45; Sopta et al, *J. Biol. Chem.* 1985, 260, 10353-60; Archambault et al., *Proc. Natl. Acad. Sci. USA* 1997, 94, 14300-5. Alternatively, the polypeptide may be tethered to the solid support through high affinity binding interactions. If the polypeptide is expressed fused to a tag, such as GST, the fusion tag can be used to anchor the polypeptide to the matrix support, for example Sepharose beads containing immobilized glutathione. Solid supports that take advantage of these tags are commercially available.

In another aspect, the support to which a polypeptide may be immobilized is a soluble support, which may facilitate certain steps performed in the methods of the present invention. For example, the soluble support may be soluble in the conditions employed to create a binding interaction between a target and the polypeptide, and then used under conditions in which it is a solid for elution of the proteins or other biological materials that bind to a polypeptide.

The concentration of the coupled polypeptide may have an affect on the sensitivity of the method. In certain embodiments, to detect interactions most efficiently, the concentration of the polypeptide bound to the matrix should be at least 10-fold higher than the K_d of the interaction. Thus, the concentration of the polypeptide bound to the matrix should be highest for the detection of the weakest protein-protein interactions. However, if the concentration of the immobilized polypeptide is not as high as may be ideal, it may still be possible to observe protein-protein interactions of interest by, for example, increasing the concentration of the polypeptide or other moiety that interacts with the coupled polypeptide. The level of detection will of course vary with each different polypeptide, interactor, conditions of the assay, etc. In certain instances, the interacting protein binds to the polypeptide with a K_d of about 10^{-5} M to about 10^{-8} M or 10^{-10} M.

In another aspect, the coupling may be done at various ratios of the polypeptide to the resin. An upper limit of the protein : resin ratio may be determined by the isoelectric

point and the ionic nature of the protein, although it may be possible to achieve higher polypeptide concentrations by use of various methods.

In certain embodiments, several concentrations of the polypeptide immobilized on a solid or soluble support may be used. One advantage of using multiple concentrations, although not a requirement, is that one may be able to obtain an estimate for the strength of the protein-protein interaction that is observed in the affinity chromatography experiment. Another advantage of using multiple concentrations is that a binding curve which has the proper shape may indicate that the interaction that is observed is biologically important rather than a spurious interaction with denatured protein.

In one example of such an embodiment, a series of columns may be prepared with varying concentrations of polypeptide (mg polypeptide/ml resin volume). The number of columns employed may be between 2 to 8, 10, 12, 15, 25 or more, each with a different concentration of attached polypeptide. Larger numbers of columns may be used if appropriate for the polypeptide being examined, and multiple columns may be used with the same concentration as any methods may require. In certain embodiments, 4 to 6 columns are prepared with varying concentrations of polypeptide. In another aspect of this embodiment, two control columns may be prepared: one that contains no polypeptide and a second that contains the highest concentration of polypeptide but is not treated with extract. After elution of the columns and separation of the eluent components (by one of the methods described below), it may be possible to distinguish the interacting proteins (if any) from the non-specific bound proteins as follows. The concentration of the interacting proteins, as determined by the intensity of the band on the gel, will increase proportionally to the increase in polypeptide concentration but will be missing from the second control column. This allows for the identification of unknown interacting proteins.

The method of the invention may be used for small-scale analysis. A variety of column sizes, types, and geometries may be used. In addition, other vessel shapes and sizes having a smaller scale than is usually found in laboratory experiments may be used as well, including a plurality of wells in a plate. For high throughput analysis, it is advantageous to use small volumes, from about 20, 30, 50, 80 or 100 μ l. Larger or small volumes may be used, as necessary, and it may be possible to achieve high throughput analysis using them. The entire affinity chromatography procedure may be automated by assembling the micro-columns into an array (e.g. with 96 micro-column arrays).

A variety of materials may be used as the source of potential interacting proteins. In one embodiment, a cellular extract or extracellular fluid may be used. The choice of starting material for the extract may be based upon the cell or tissue type or type of fluid that would be expected to contain proteins that interact with the target protein. Micro-organisms or other organisms are grown in a medium that is appropriate for that organism and can be grown in specific conditions to promote the expression of proteins that may interact with the target protein. Exemplary starting material that may be used to make a suitable extract are: 1) one or more types of tissue derived from an animal, plant, or other multi-cellular organism, 2) cells grown in tissue culture that were derived from an animal or human, plant or other source, 3) micro-organisms grown in suspension or non-suspension cultures, 4) virus-infected cells, 5) purified organelles (including, but not restricted to nuclei, mitochondria, membranes, Golgi, endoplasmic reticulum, lysosomes, or peroxisomes) prepared by differential centrifugation or another procedure from animal, plant or other kinds of eukaryotic cells, 6) serum or other bodily fluids including, but not limited to, blood, urine, semen, synovial fluid, cerebrospinal fluid, amniotic fluid, lymphatic fluid or interstitial fluid. In other embodiments, a total cell extract may not be the optimal source of interacting proteins. For example, if the ligand is known to act in the nucleus, a nuclear extract can provide a 10-fold enrichment of proteins that are likely to interact with the ligand. In addition, proteins that are present in the extract in low concentrations may be enriched using another chromatographic method to fractionate the extract before screening various pools for an interacting protein.

Extracts are prepared by methods known to those of skill in the art. The extracts may be prepared at a low temperature (e.g., 4°C) in order to retard denaturation or degradation of proteins in the extract. The pH of the extract may be adjusted to be appropriate for the body fluid or tissue, cellular, or organellar source that is used for the procedure (e.g. pH 7-8 for cytosolic extracts from mammals, but low pH for lysosomal extracts). The concentration of chaotropic or non-chaotropic salts in the extracting solution may be adjusted so as to extract the appropriate sets of proteins for the procedure. Glycerol may be added to the extract, as it aids in maintaining the stability of many proteins and also reduces background non-specific binding. Both the lysis buffer and column buffer may contain protease inhibitors to minimize proteolytic degradation of proteins in the extract and to protect the polypeptide. Appropriate co-factors that could potentially interact with the interacting proteins may be added to the extracting solution. One or more nucleases or

another reagent may be added to the extract, if appropriate, to prevent protein-protein interactions that are mediated by nucleic acids. Appropriate detergents or other agents may be added to the solution, if desired, to extract membrane proteins from the cells or tissue. A reducing agent (e.g. dithiothreitol or 2-mercaptoethanol or glutathione or other agent) may be added. Trace metals or a chelating agent may be added, if desired, to the extracting solution.

Usually, the extract is centrifuged in a centrifuge or ultracentrifuge or filtered to provide a clarified supernatant solution. This supernatant solution may be dialyzed using dialysis tubing, or another kind of device that is standard in the art, against a solution that is similar to, but may not be identical with, the solution that was used to make the extract. The extract is clarified by centrifugation or filtration again immediately prior to its use in affinity chromatography.

In some cases, the crude lysate will contain small molecules that can interfere with the affinity chromatography. This can be remedied by precipitating proteins with ammonium sulfate, centrifugation of the precipitate, and re-suspending the proteins in the affinity column buffer followed by dialysis. An additional centrifugation of the sample may be needed to remove any particulate matter prior to application to the affinity columns.

The amount of cell extract applied to the column may be important for any embodiment. If too little extract is applied to the column and the interacting protein is present at low concentration, the level of interacting protein retained by the column may be difficult to detect. Conversely, if too much extract is applied to the column, protein may precipitate on the column or competition by abundant interacting proteins for the limited amount of protein ligand may result in a difficulty in detecting minor species.

The columns functionalized with a polypeptide of the invention are loaded with protein extract from an appropriate source that has been dialyzed against a buffer that is consistent with the nature of the expected interaction. The pH, salt concentrations and the presence or absence of reducing and chelating agents, trace metals, detergents, and co-factors may be adjusted according to the nature of the expected interaction. Most commonly, the pH and the ionic strength are chosen so as to be close to physiological for the source of the extract. The extract is most commonly loaded under gravity onto the columns at a flow rate of about 4-6 column volumes per hour, but this flow rate can be adjusted for particular circumstances in an automated procedure.

The volume of the extract that is loaded on the columns can be varied but is most commonly equivalent to about 5 to 10 column volumes. When large volumes of extract are loaded on the columns, there is often an improvement in the signal-to-noise ratio because more protein from the extract is available to bind to the protein ligand, whereas the background binding of proteins from the extract to the solid support saturates with low amounts of extract.

A control column may be included that contains the highest concentration of protein ligand, but buffer rather than extract is loaded onto this column. The elutions (eluates) from this column will contain polypeptide that failed to be attached to the column in a covalent manner, but no proteins that are derived from the extract.

The columns may be washed with a buffer appropriate to the nature of the interaction being analyzed, usually, but not necessarily, the same as the loading buffer. An elution buffer with an appropriate pH, glycerol, and the presence or absence of reducing agent, chelating agent, cofactors, and detergents are all important considerations. The columns may be washed with anywhere from about 5 to 20 column volumes of each wash buffer to eliminate unbound proteins from the natural extract. The flow rate of the wash is usually adjusted to about 4 to 6 column volumes per hour by using gravity or an automated procedure, but other flow rates are possible in specific circumstances.

In order to elute the proteins that have been retained by the column, the interactions between the extract proteins and the column ligand should be disrupted. This is performed by eluting the column with a solution of salt or detergent. Retention of activity by the eluted proteins may require the presence of glycerol and a buffer of appropriate pH, as well as proper choices of ionic strength and the presence or absence of appropriate reducing agent, chelating agent, trace metals, cofactors, detergents, chaotropic agents, and other reagents. If physical identification of the bound proteins is the objective, the elution may be performed sequentially, first with buffer of high ionic strength and then with buffer containing a protein denaturant, most commonly, but not restricted to sodium dodecyl sulfate (SDS), urea, or guanidine hydrochloride. In certain instances, the column is eluted with a protein denaturant, particularly SDS, for example as a 1% SDS solution. Using only the SDS wash, and omitting the salt wash, may result in SDS-gels that have higher resolution (sharper bands with less smearing). Also, using only the SDS wash results in half as many samples to analyze. The volume of the eluting solution may be varied but is normally about 2 to 4 column volumes. For 20 ml columns, the flow rate of the eluting

procedures are most commonly about 4 to 6 column volumes per hour, under gravity, but can be varied in an automated procedure.

The proteins from the extract that were bound to and are eluted from the affinity columns may be most easily resolved for identification by an electrophoresis procedure, but this procedure may be modified, replaced by another suitable method, or omitted. Any of the denaturing or non-denaturing electrophoresis procedures that are standard in the art may be used for this purpose, including SDS-PAGE, gradient gels, capillary electrophoresis, and two-dimensional gels with isoelectric focusing in the first dimension and SDS-PAGE in the second. Typically, the individual components in the column eluent are separated by polyacrylamide gel electrophoresis.

After electrophoresis, protein bands or spots may be visualized using any number of methods known to those of skill in the art, including staining techniques such as Coomassie blue or silver staining, or some other agent that is standard in the art. Alternatively, autoradiography can be used for visualizing proteins isolated from organisms cultured on media containing a radioactive label, for example $^{35}\text{SO}_4^{2-}$ or $^{35}\text{[S]}$ methionine, that is incorporated into the proteins. The use of radioactively labeled extract allows a distinction to be made between extract proteins that were retained by the column and proteolytic fragments of the ligand that may be released from the column.

Protein bands that are derived from the extract (i.e. it did not elute from the control column that was not loaded with protein from the extract) and bound to an experimental column that contained polypeptide covalently attached to the solid support, and did not bind to a control column that did not contain any polypeptide, may be excised from the stained electrophoretic gel and further characterized.

To identify the protein interactor by mass spectrometry, it may be desirable to reduce the disulfide bonds of the protein followed by alkylation of the free thiols prior to digestion of the protein with protease. The reduction may be performed by treatment of the gel slice with a reducing agent, for example with dithiothreitol, whereupon, the protein is alkylated by treating the gel slice with a suitable alkylating agent, for example iodoacetamide.

Prior to analysis by mass spectrometry, the protein may be chemically or enzymatically digested. The protein sample in the gel slice may be subjected to *in-gel* digestion. Shevchenko A. et al., Mass Spectrometric Sequencing of Proteins from Silver

Stained Polyacrylamide Gels. Analytical Chemistry 1996, 58, 850-858. One method of digestion is by treatment with the enzyme trypsin. The resulting peptides are extracted from the gel slice into a buffer.

The peptide fragments may be purified, for example by use of chromatography. A solid support that differentially binds the peptides and not the other compounds derived from the gel slice, the protease reaction or the peptide extract may be used. The peptides may be eluted from the solid support into a small volume of a solution that is compatible with mass spectrometry (e.g. 50% acetonitrile/0.1% trifluoroacetic acid).

The preparation of a protein sample from a gel slice that is suitable for mass spectrometry may also be done by an automated procedure.

Peptide samples derived from gel slices may be analyzed by any one of a variety of techniques in mass spectrometry as further described above. This technique may be used to assign function to an unknown protein based upon the known function of the interacting protein in the same or a homologous/orthologous organism.

Eluates from the affinity chromatography columns may also be analyzed directly without resolution by electrophoretic methods, by proteolytic digestion with a protease in solution, followed by applying the proteolytic digestion products to a reverse phase column and eluting the peptides from the column.

In yet another embodiment, proteins that interact with a polypeptide of the invention may be identified using an interaction trap assay (see also, U.S. Patent NO: 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J Biol Chem* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; and Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696).

In another embodiment, a method of the present invention makes use of chimeric genes which express hybrid proteins. To illustrate, a first hybrid gene comprises the coding sequence for a DNA-binding domain of a transcriptional activator fused in frame to the coding sequence for a "bait" protein, e.g., a polypeptide of the invention of sufficient length to bind to a potential interacting protein. The second hybrid protein encodes a transcriptional activation domain fused in frame to a gene encoding a "fish" protein, e.g., a potential interacting protein of sufficient length to interact with a polypeptide of the invention portion of the bait fusion protein. If the bait and fish proteins are able to interact, e.g., form a protein-protein interaction, they bring into close proximity the two domains of

the transcriptional activator. This proximity causes transcription of a reporter gene which is operably linked to a transcriptional regulatory site responsive to the transcriptional activator, and expression of the reporter gene can be detected and used to score for the interaction of the bait and fish proteins.

In accordance with the present invention, the method includes providing a host cell, typically a yeast cell, e.g., *Kluyverei lactis*, *Schizosaccharomyces pombe*, *Ustilago maydis*, *Saccharomyces cerevisiae*, *Neurospora crassa*, *Aspergillus niger*, *Aspergillus nidulans*, *Pichia pastoris*, *Candida tropicalis*, and *Hansenula polymorpha*, though most preferably *S. cerevisiae* or *S. pombe*. The host cell contains a reporter gene having a binding site for the DNA-binding domain of a transcriptional activator used in the bait protein, such that the reporter gene expresses a detectable gene product when the gene is transcriptionally activated. The first chimeric gene may be present in a chromosome of the host cell, or as part of an expression vector.

The host cell also contains a first chimeric gene which is capable of being expressed in the host cell. The gene encodes a chimeric protein, which comprises (a) a DNA-binding domain that recognizes the responsive element on the reporter gene in the host cell, and (b) a bait protein (e.g., a polypeptide of the invention).

A second chimeric gene is also provided which is capable of being expressed in the host cell, and encodes the "fish" fusion protein. In one embodiment, both the first and the second chimeric genes are introduced into the host cell in the form of plasmids. Preferably, however, the first chimeric gene is present in a chromosome of the host cell and the second chimeric gene is introduced into the host cell as part of a plasmid.

The DNA-binding domain of the first hybrid protein and the transcriptional activation domain of the second hybrid protein may be derived from transcriptional activators having separable DNA-binding and transcriptional activation domains. For instance, these separate DNA-binding and transcriptional activation domains are known to be found in the yeast GAL4 protein, and are known to be found in the yeast GCN4 and ADR1 proteins. Many other proteins involved in transcription also have separable binding and transcriptional activation domains which make them useful for the present invention, and include, for example, the LexA and VP16 proteins. It will be understood that other (substantially) transcriptionally-inert DNA-binding domains may be used in the subject constructs; such as domains of ACE1, λ CI, lac repressor, jun or fos. In another

embodiment, the DNA-binding domain and the transcriptional activation domain may be from different proteins. The use of a LexA DNA binding domain provides certain advantages. For example, in yeast, the LexA moiety contains no activation function and has no known effect on transcription of yeast genes. In addition, use of LexA allows control over the sensitivity of the assay to the level of interaction (see, for example, the Brent *et al.* PCT publication WO94/10300).

In certain embodiments, any enzymatic activity associated with the bait or fish proteins is inactivated, e.g., dominant negative or other mutants of a protein-protein interaction component can be used.

Continuing with the illustrative example, a polypeptide of the invention-mediated interaction, if any, between the bait and fish fusion proteins in the host cell, causes the activation domain to activate transcription of the reporter gene. The method is carried out by introducing the first chimeric gene and the second chimeric gene into the host cell, and subjecting that cell to conditions under which the bait and fish fusion proteins are expressed in sufficient quantity for the reporter gene to be activated. The formation of a protein complex containing a polypeptide of the invention results in a detectable signal produced by the expression of the reporter gene.

In still further embodiments, the protein-protein interaction of interest is generated in whole cells, taking advantage of cell culture techniques to support the subject assay. For example, the protein-protein interaction of interest can be constituted in a prokaryotic or eukaryotic cell culture system. Advantages to generating the protein complex in an intact cell includes the ability to screen for inhibitors of the level or activity of the complex which are functional in an environment more closely approximating that which therapeutic use of the inhibitor would require, including the ability of the agent to gain entry into the cell. Furthermore, certain of the *in vivo* embodiments of the assay are amenable to high throughput analysis of candidate agents.

The components of the protein complex comprising a polypeptide of the invention can be endogenous to the cell selected to support the assay. Alternatively, some or all of the components can be derived from exogenous sources. For instance, fusion proteins can be introduced into the cell by recombinant techniques (such as through the use of an expression vector), as well as by microinjecting the fusion protein itself or mRNA encoding the fusion protein. Moreover, in the whole cell embodiments of the subject assay, the

reporter gene construct can provide, upon expression, a selectable marker. Such embodiments of the subject assay are particularly amenable to high through-put analysis in that proliferation of the cell can provide a simple measure of the protein-protein interaction.

The amount of transcription from the reporter gene may be measured using any method known to those of skill in the art to be suitable. For example, specific mRNA expression may be detected using Northern blots or specific protein product may be identified by a characteristic stain, western blots or an intrinsic activity. In certain embodiments, the product of the reporter gene is detected by an intrinsic activity associated with that product. For instance, the reporter gene may encode a gene product that, by enzymatic activity, gives rise to a detection signal based on color, fluorescence, or luminescence.

The interaction trap assay of the invention may also be used to identify test agents capable of modulating formation of a complex comprising a polypeptide of the invention. In general, the amount of expression from the reporter gene in the presence of the test compound is compared to the amount of expression in the same cell in the absence of the test compound. Alternatively, the amount of expression from the reporter gene in the presence of the test compound may be compared with the amount of transcription in a substantially identical cell that lacks a component of the protein-protein interaction involving a polypeptide of the invention.

7. Antibodies

Another aspect of the invention pertains to antibodies specifically reactive with a polypeptide of the invention. For example, by using peptides based on a polypeptide of the invention, e.g., having an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or an immunogenic fragment thereof, antisera or monoclonal antibodies may be made using standard methods. An exemplary immunogenic fragment may contain eight, ten or more consecutive amino acid residues of SEQ ID NO: 2 or SEQ ID NO: 4. Certain fragments that are predicted to be immunogenic for the subject amino acid sequences (predicted) are set forth in Table 2 contained in FIGURE 7.

The term "antibody" as used herein is intended to include fragments thereof which are also specifically reactive with a polypeptide of the invention. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the

same manner as is suitable for whole antibodies. For example, $F(ab')_2$ fragments can be generated by treating antibody with pepsin. The resulting $F(ab')_2$ fragment can be treated to reduce disulfide bridges to produce Fab' fragments. The antibody of the present invention is further intended to include bispecific and chimeric molecules, as well as single chain (scFv) antibodies. Also within the scope of the invention are trimeric antibodies, humanized antibodies, human antibodies, and single chain antibodies. All of these modified forms of antibodies as well as fragments of antibodies are intended to be included in the term "antibody".

In one aspect, the present invention contemplates a purified antibody that binds specifically to a polypeptide of the invention and which does not substantially cross-react with a protein which is less than about 80%, or less than about 90%, identical to SEQ ID NO: 2 or SEQ ID NO: 4. In another aspect, the present invention contemplates an array comprising a substrate having a plurality of address, wherein at least one of the addresses has disposed thereon a purified antibody that binds specifically to a polypeptide of the invention.

Antibodies may be elicited by methods known in the art. For example, a mammal such as a mouse, a hamster or rabbit may be immunized with an immunogenic form of a polypeptide of the invention (e.g., an antigenic fragment which is capable of eliciting an antibody response). Alternatively, immunization may occur by using a nucleic acid of the acid, which presumably *in vivo* expresses the polypeptide of the invention giving rise to the immunogenic response observed. Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. For instance, a peptidyl portion of a polypeptide of the invention may be administered in the presence of adjuvant. The progress of immunization may be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays may be used with the immunogen as antigen to assess the levels of antibodies.

Following immunization, antisera reactive with a polypeptide of the invention may be obtained and, if desired, polyclonal antibodies isolated from the serum. To produce monoclonal antibodies, antibody producing cells (lymphocytes) may be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique (originally developed by Kohler

and Milstein, (1975) *Nature*, 256: 495-497), as the human B cell hybridoma technique (Kozbar et al., (1983) *Immunology Today*, 4: 72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with the polypeptides of the invention and the monoclonal antibodies isolated.

Antibodies directed against the polypeptides of the invention can be used to selectively block the action of the polypeptides of the invention. Antibodies against a polypeptide of the invention may be employed to treat infections, particularly bacterial infections and diseases. For example, the present invention contemplates a method for treating a subject suffering from a *P. aeruginosa* related disease or disorder, comprising administering to an animal having the condition a therapeutically effective amount of a purified antibody that binds specifically to a polypeptide of the invention. In another example, the present invention contemplates a method for inhibiting SEQ ID NO: 2 or SEQ ID NO: 4 dependent growth or infectivity of *P. aeruginosa*, comprising contacting *P. aeruginosa* with a purified antibody that binds specifically to a polypeptide of the invention.

In one embodiment, antibodies reactive with a polypeptide of the invention are used in the immunological screening of cDNA libraries constructed in expression vectors, such as λ gt11, λ gt18-23, λ ZAP, and λ ORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance, λ gt11 will produce fusion proteins whose amino termini consist of β -galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of a polypeptide of the invention can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from phage infected bacterial plates with an antibody specific for a polypeptide of the invention. Phage scored by this assay can then be isolated from the infected plate. Thus, homologs of a polypeptide of the invention can be detected and cloned from other sources.

Antibodies may be employed to isolate or to identify clones expressing the polypeptides to purify the polypeptides by affinity chromatography.

In other embodiments, the polypeptides of the invention may be modified so as to increase their immunogenicity. For example, a polypeptide, such as an antigenically or

immunologically equivalent derivative, may be associated, for example by conjugation, with an immunogenic carrier protein for example bovine serum albumin (BSA) or keyhole limpet haemocyanin (KLH). Alternatively a multiple antigenic peptide comprising multiple copies of the protein or polypeptide, or an antigenically or immunologically equivalent polypeptide thereof may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

In other embodiments, the antibodies of the invention, or variants thereof, are modified to make them less immunogenic when administered to a subject. For example, if the subject is human, the antibody may be "humanized"; where the complementarity determining region(s) of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones, P. et al. (1986), *Nature* 321, 522-525 or Tempest et al. (1991) *Biotechnology* 9, 266-273. Also, transgenic mice, or other mammals, may be used to express humanized antibodies. Such humanization may be partial or complete.

The use of a nucleic acid of the invention in genetic immunization may employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff et al., *Hum Mol Genet* 1992, 1:363, Manthorpe et al., *Hum. Gene Ther.* 1993:4, 419), delivery of DNA complexed with specific protein carriers (Wu et al., *J Biol Chem.* 1989: 264,16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, *PNAS USA*, 1986:83,9551), encapsulation of DNA in various forms of liposomes (Kaneda et al., *Science* 1989:243,375), particle bombardment (Tang et al., *Nature* 1992, 356:152, Eisenbraun et al., *DNA Cell Biol* 1993, 12:791) and *in vivo* infection using cloned retroviral vectors (Seeger et al., *PNAS USA* 1984:81,5849).

8. Diagnostic Assays

The invention further provides a method for detecting the presence of *P. aeruginosa* in a biological sample. Detection of *P. aeruginosa* in a subject, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of a *P. aeruginosa* related disease or disorder. In general, the method involves contacting the biological sample with a compound or an agent capable of detecting a polypeptide of the invention or a nucleic acid of the invention. The term "biological sample" when used in reference to a

diagnostic assay is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject.

The detection method of the invention may be used to detect the presence of *P. aeruginosa* in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of a nucleic acid of the invention include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of polypeptides of the invention include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations, immunofluorescence, radioimmunoassays and competitive binding assays. Alternatively, polypeptides of the invention can be detected *in vivo* in a subject by introducing into the subject a labeled antibody specific for a polypeptide of the invention. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques. It may be possible to use all of the diagnostic methods disclosed herein for pathogens in addition to *P. aeruginosa*.

Nucleic acids for diagnosis may be obtained from an infected individual's cells and tissues, such as bone, blood, muscle, cartilage, and skin. Nucleic acids, e.g., DNA and RNA, may be used directly for detection or may be amplified, e.g., enzymatically by using PCR or other amplification technique, prior to analysis. Using amplification, characterization of the species and strain of prokaryote present in an individual, may be made by an analysis of the genotype of the prokaryote gene. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the genotype of a reference sequence. Point mutations can be identified by hybridizing a nucleic acid, e.g., amplified DNA, to a nucleic acid of the invention, which nucleic acid may be labeled. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in the electrophoretic mobility of the DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. See, e.g. Myers et al., Science, 230: 1242 (1985). Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or a chemical cleavage method. See, e.g., Cotton et al., Proc. Natl. Acad. Sci., USA, 85: 4397-4401 (1985).

Agents for detecting a nucleic acid of the invention, e.g., comprising the sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 3, include labeled or labelable nucleic acid probes capable of hybridizing to a nucleic acid of the invention. The nucleic acid probe can

comprise, for example, the full length sequence of a nucleic acid of the invention, or an equivalent thereof, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to SEQ ID NO: 1 or SEQ ID NO: 3, or the complement thereof. Agents for detecting a polypeptide of the invention, e.g., comprising an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, include labeled or labelable antibodies capable of binding to a polypeptide of the invention. Antibodies may be polyclonal, or alternatively, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')₂) can be used. Labeling the probe or antibody also encompasses direct labeling of the probe or antibody by coupling (e.g., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin.

In certain embodiments, detection of a nucleic acid of the invention in a biological sample involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) Science 241:1077-1080; and Nakazawa et al. (1994) PNAS 91:360-364), the latter of which can be particularly useful for distinguishing between orthologs of polynucleotides of the invention (see Abravaya et al. (1995) Nucleic Acids Res. 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a nucleic acid of the invention under conditions such that hybridization and amplification of the polynucleotide (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample.

In one aspect, the present invention contemplates a method for detecting the presence of *P. aeruginosa* in a sample, the method comprising: (a) providing a sample to be tested for the presence of *P. aeruginosa*; (b) contacting the sample with an antibody reactive against eight consecutive amino acid residues of SEQ ID NO: 2 or SEQ ID NO: 4 under conditions which permit association between the antibody and its ligand; and

(c) detecting interaction of the antibody with its ligand, thereby detecting the presence of *P. aeruginosa* in the sample.

In another aspect, the present invention contemplates a method for detecting the presence of *P. aeruginosa* in a sample, the method comprising: (a) providing a sample to be tested for the presence of *P. aeruginosa*; (b) contacting the sample with an antibody that binds specifically to a polypeptide of the invention under conditions which permit association between the antibody and its ligand; and (c) detecting interaction of the antibody with its ligand, thereby detecting the presence of *P. aeruginosa* in the sample.

In yet another example, the present invention contemplates a method for diagnosing a patient suffering from a *P. aeruginosa* related disease or disorder, comprising: (a) obtaining a biological sample from a patient; (b) detecting the presence or absence of a polypeptide of the invention, or a nucleic acid encoding a polypeptide of the invention, in the sample; and (c) diagnosing a patient suffering from a *P. aeruginosa* related disease or disorder based on the presence of a polypeptide of the invention, or a nucleic acid encoding a polypeptide of the invention, in the patient sample.

The diagnostic assays of the invention may also be used to monitor the effectiveness of an anti-*P. aeruginosa* treatment in an individual suffering from an *P. aeruginosa* related disease or disorder. For example, the presence and/or amount of a nucleic acid of the invention or a polypeptide of the invention can be detected in an individual suffering from an *P. aeruginosa* related disease or disorder before and after treatment with anti-*P. aeruginosa* therapeutic agent. Any change in the level of a polynucleotide or polypeptide of the invention after treatment of the individual with the therapeutic agent can provide information about the effectiveness of the treatment course. In particular, no change, or a decrease, in the level of a polynucleotide or polypeptide of the invention present in the biological sample will indicate that the therapeutic is successfully combating the *P. aeruginosa* related disease or disorder.

The invention also encompasses kits for detecting the presence of *P. aeruginosa* in a biological sample. For example, the kit can comprise a labeled or labelable compound or agent capable of detecting a polynucleotide or polypeptide of the invention in a biological sample; means for determining the amount of *P. aeruginosa* in the sample; and means for comparing the amount of *P. aeruginosa* in the sample with a standard. The compound or

agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect a polynucleotide or polypeptide of the invention.

9. Drug Discovery

Modulators to polypeptides of the invention and other structurally related molecules, and complexes containing the same, may be identified and developed as set forth below and otherwise using techniques and methods known to those of skill in the art. The modulators of the invention may be employed, for instance, to inhibit and treat *P. aeruginosa* associated diseases or conditions, such as osteomyelitis, otitis externa, conjunctivitis, keratitis, endophthalmitis, alveolar necrosis, vascular invasion, bacteremia, and burn infection.

A variety of methods for inhibiting the growth or infectivity of *P. aeruginosa* are contemplated by the present invention. For example, exemplary methods involve contacting *P. aeruginosa* with a polypeptide of the invention that modulates the same or another polypeptide from such pathogen, a nucleic acid encoding such polypeptide of the invention, or a compound thought or shown to be effective against such pathogen.

For example, in one aspect, the present invention contemplates a method for treating a patient suffering from an infection of *P. aeruginosa*, comprising administering to the patient an amount of a SEQ ID NO: 2 or SEQ ID NO: 4 inhibitor effective to inhibit the expression and/or activity of a polypeptide of the invention. In certain instances, the animal is a human or a livestock animal such as a cow, pig, goat or sheep. The present invention further contemplates a method for treating a subject suffering from a *P. aeruginosa* related disease or disorder, comprising administering to an animal having the condition a therapeutically effective amount of a molecule identified using one of the methods of the present invention.

The present invention contemplates making any molecule that is shown to modulate the activity of a polypeptide of the invention.

In another embodiment, inhibitors, modulators of the subject polypeptides, or biological complexes containing them, may be used in the manufacture of a medicament for any number of uses, including, for example, treating any disease or other treatable condition of a patient (including humans and animals), and particularly a disease caused by *P.*

aeruginosa, such as, for example, one of the following: osteomyelitis, otitis externa, conjunctivitis, keratitis, endophthalmitis, alveolar necrosis, vascular invasion, bacteremia, and burn infection..

(a) Drug Design

A number of techniques can be used to screen, identify, select and design chemical entities capable of associating with polypeptides of the invention, structurally homologous molecules, and other molecules. Knowledge of the structure for a polypeptide of the invention, determined in accordance with the methods described herein, permits the design and/or identification of molecules and/or other modulators which have a shape complementary to the conformation of a polypeptide of the invention, or more particularly, a druggable region thereof. It is understood that such techniques and methods may use, in addition to the exact structural coordinates and other information for a polypeptide of the invention, structural equivalents thereof described above (including, for example, those structural coordinates that are derived from the structural coordinates of amino acids contained in a druggable region as described above).

The term "chemical entity," as used herein, refers to chemical compounds, complexes of two or more chemical compounds, and fragments of such compounds or complexes. In certain instances, it is desirable to use chemical entities exhibiting a wide range of structural and functional diversity, such as compounds exhibiting different shapes (e.g., flat aromatic rings(s), puckered aliphatic rings(s), straight and branched chain aliphatics with single, double, or triple bonds) and diverse functional groups (e.g., carboxylic acids, esters, ethers, amines, aldehydes, ketones, and various heterocyclic rings).

In one aspect, the method of drug design generally includes computationally evaluating the potential of a selected chemical entity to associate with any of the molecules or complexes of the present invention (or portions thereof). For example, this method may include the steps of (a) employing computational means to perform a fitting operation between the selected chemical entity and a druggable region of the molecule or complex; and (b) analyzing the results of said fitting operation to quantify the association between the chemical entity and the druggable region.

A chemical entity may be examined either through visual inspection or through the use of computer modeling using a docking program such as GRAM, DOCK, or AUTODOCK (Dunbrack et al., *Folding & Design*, 2:27-42 (1997)). This procedure can

include computer fitting of chemical entities to a target to ascertain how well the shape and the chemical structure of each chemical entity will complement or interfere with the structure of the subject polypeptide (Bugg et al., *Scientific American*, Dec.: 92-98 (1993); West et al., *TIPS*, 16:67-74 (1995)). Computer programs may also be employed to estimate the attraction, repulsion, and steric hindrance of the chemical entity to a druggable region, for example. Generally, the tighter the fit (e.g., the lower the steric hindrance, and/or the greater the attractive force) the more potent the chemical entity will be because these properties are consistent with a tighter binding constant. Furthermore, the more specificity in the design of a chemical entity the more likely that the chemical entity will not interfere with related proteins, which may minimize potential side-effects due to unwanted interactions.

A variety of computational methods for molecular design, in which the steric and electronic properties of druggable regions are used to guide the design of chemical entities, are known: Cohen et al. (1990) *J. Med. Cam.* 33: 883-894; Kuntz et al. (1982) *J. Mol. Biol.* 161: 269-288; DesJarlais (1988) *J. Med. Cam.* 31: 722-729; Bartlett et al. (1989) *Spec. Publ., Roy. Soc. Chem.* 78: 182-196; Goodford et al. (1985) *J. Med. Cam.* 28: 849-857; and DesJarlais et al. *J. Med. Cam.* 29: 2149-2153. Directed methods generally fall into two categories: (1) design by analogy in which 3-D structures of known chemical entities (such as from a crystallographic database) are docked to the druggable region and scored for goodness-of-fit; and (2) *de novo* design, in which the chemical entity is constructed piece-wise in the druggable region. The chemical entity may be screened as part of a library or a database of molecules. Databases which may be used include ACD (Molecular Designs Limited), NCI (National Cancer Institute), CCDC (Cambridge Crystallographic Data Center), CAST (Chemical Abstract Service), Derwent (Derwent Information Limited), Maybridge (Maybridge Chemical Company Ltd), Aldrich (Aldrich Chemical Company), DOCK (University of California in San Francisco), and the Directory of Natural Products (Chapman & Hall). Computer programs such as CONCORD (Tripos Associates) or DB-Converter (Molecular Simulations Limited) can be used to convert a data set represented in two dimensions to one represented in three dimensions.

Chemical entities may be tested for their capacity to fit spatially with a druggable region or other portion of a target protein. As used herein, the term "fits spatially" means that the three-dimensional structure of the chemical entity is accommodated geometrically by a druggable region. A favorable geometric fit occurs when the surface area of the

chemical entity is in close proximity with the surface area of the druggable region without forming unfavorable interactions. A favorable complementary interaction occurs where the chemical entity interacts by hydrophobic, aromatic, ionic, dipolar, or hydrogen donating and accepting forces. Unfavorable interactions may be steric hindrance between atoms in the chemical entity and atoms in the druggable region.

If a model of the present invention is a computer model, the chemical entities may be positioned in a druggable region through computational docking. If, on the other hand, the model of the present invention is a structural model, the chemical entities may be positioned in the druggable region by, for example, manual docking. As used herein the term "docking" refers to a process of placing a chemical entity in close proximity with a druggable region, or a process of finding low energy conformations of a chemical entity/druggable region complex.

In an illustrative embodiment, the design of potential modulator begins from the general perspective of shape complimentary for the druggable region of a polypeptide of the invention, and a search algorithm is employed which is capable of scanning a database of small molecules of known three-dimensional structure for chemical entities which fit geometrically with the target druggable region. Most algorithms of this type provide a method for finding a wide assortment of chemical entities that are complementary to the shape of a druggable region of the subject polypeptide. Each of a set of chemical entities from a particular data-base, such as the Cambridge Crystallographic Data Bank (CCDB) (Allen et al. (1973) *J. Chem. Doc.* 13: 119), is individually docked to the druggable region of a polypeptide of the invention in a number of geometrically permissible orientations with use of a docking algorithm. In certain embodiments, a set of computer algorithms called DOCK, can be used to characterize the shape of invaginations and grooves that form the active sites and recognition surfaces of the druggable region (Kuntz et al. (1982) *J. Mol. Biol.* 161: 269-288). The program can also search a database of small molecules for templates whose shapes are complementary to particular binding sites of a polypeptide of the invention (DesJarlais et al. (1988) *J Med Chem* 31: 722-729).

The orientations are evaluated for goodness-of-fit and the best are kept for further examination using molecular mechanics programs, such as AMBER or CHARMM. Such algorithms have previously proven successful in finding a variety of chemical entities that are complementary in shape to a druggable region.

Goodford (1985, *J Med Chem* 28:849-857) and Boobbyer et al. (1989, *J Med Chem* 32:1083-1094) have produced a computer program (GRID) which seeks to determine regions of high affinity for different chemical groups (termed probes) of the druggable region. GRID hence provides a tool for suggesting modifications to known chemical entities that might enhance binding. It may be anticipated that some of the sites discerned by GRID as regions of high affinity correspond to "pharmacophoric patterns" determined inferentially from a series of known ligands. As used herein, a "pharmacophoric pattern" is a geometric arrangement of features of chemical entities that is believed to be important for binding. Attempts have been made to use pharmacophoric patterns as a search screen for novel ligands (Jakes et al. (1987) *J Mol Graph* 5:41-48; Brint et al. (1987) *J Mol Graph* 5:49-56; Jakes et al. (1986) *J Mol Graph* 4:12-20).

Yet a further embodiment of the present invention utilizes a computer algorithm such as CLIX which searches such databases as CCDB for chemical entities which can be oriented with the druggable region in a way that is both sterically acceptable and has a high likelihood of achieving favorable chemical interactions between the chemical entity and the surrounding amino acid residues. The method is based on characterizing the region in terms of an ensemble of favorable binding positions for different chemical groups and then searching for orientations of the chemical entities that cause maximum spatial coincidence of individual candidate chemical groups with members of the ensemble. The algorithmic details of CLIX is described in Lawrence et al. (1992) *Proteins* 12:31-41.

In this way, the efficiency with which a chemical entity may bind to or interfere with a druggable region may be tested and optimized by computational evaluation. For example, for a favorable association with a druggable region, a chemical entity must preferably demonstrate a relatively small difference in energy between its bound and free states (i.e., a small deformation energy of binding). Thus, certain, more desirable chemical entities will be designed with a deformation energy of binding of not greater than about 10 kcal/mole, and more preferably, not greater than 7 kcal/mole. Chemical entities may interact with a druggable region in more than one conformation that is similar in overall binding energy. In those cases, the deformation energy of binding is taken to be the difference between the energy of the free entity and the average energy of the conformations observed when the chemical entity binds to the target.

In this way, the present invention provides computer-assisted methods for identifying or designing a potential modulator of the activity of a polypeptide of the

invention including: supplying a computer modeling application with a set of structure coordinates of a molecule or complex, the molecule or complex including at least a portion of a druggable region from a polypeptide of the invention; supplying the computer modeling application with a set of structure coordinates of a chemical entity; and determining whether the chemical entity is expected to bind to the molecule or complex, wherein binding to the molecule or complex is indicative of potential modulation of the activity of a polypeptide of the invention.

In another aspect, the present invention provides a computer-assisted method for identifying or designing a potential modulator to a polypeptide of the invention, supplying a computer modeling application with a set of structure coordinates of a molecule or complex, the molecule or complex including at least a portion of a druggable region of a polypeptide of the invention; supplying the computer modeling application with a set of structure coordinates for a chemical entity; evaluating the potential binding interactions between the chemical entity and active site of the molecule or molecular complex; structurally modifying the chemical entity to yield a set of structure coordinates for a modified chemical entity, and determining whether the modified chemical entity is expected to bind to the molecule or complex, wherein binding to the molecule or complex is indicative of potential modulation of the polypeptide of the invention.

In one embodiment, a potential modulator can be obtained by screening a peptide library (Scott and Smith, *Science*, 249:386-390 (1990); Cwirla et al., *Proc. Natl. Acad. Sci.*, 87:6378-6382 (1990); Devlin et al., *Science*, 249:404-406 (1990)). A potential modulator selected in this manner could then be systematically modified by computer modeling programs until one or more promising potential drugs are identified. Such analysis has been shown to be effective in the development of HIV protease inhibitors (Lam et al., *Science* 263:380-384 (1994); Wlodawer et al., *Ann. Rev. Biochem.* 62:543-585 (1993); Appelt, *Perspectives in Drug Discovery and Design* 1:23-48 (1993); Erickson, *Perspectives in Drug Discovery and Design* 1:109-128 (1993)). Alternatively a potential modulator may be selected from a library of chemicals such as those that can be licensed from third parties, such as chemical and pharmaceutical companies. A third alternative is to synthesize the potential modulator *de novo*.

For example, in certain embodiments, the present invention provides a method for making a potential modulator for a polypeptide of the invention, the method including synthesizing a chemical entity or a molecule containing the chemical entity to yield a

potential modulator of a polypeptide of the invention, the chemical entity having been identified during a computer-assisted process including supplying a computer modeling application with a set of structure coordinates of a molecule or complex, the molecule or complex including at least one druggable region from a polypeptide of the invention; supplying the computer modeling application with a set of structure coordinates of a chemical entity; and determining whether the chemical entity is expected to bind to the molecule or complex at the active site, wherein binding to the molecule or complex is indicative of potential modulation. This method may further include the steps of evaluating the potential binding interactions between the chemical entity and the active site of the molecule or molecular complex and structurally modifying the chemical entity to yield a set of structure coordinates for a modified chemical entity, which steps may be repeated one or more times.

Once a potential modulator is identified, it can then be tested in any standard assay for the macromolecule depending of course on the macromolecule, including in high throughput assays. Further refinements to the structure of the modulator will generally be necessary and can be made by the successive iterations of any and/or all of the steps provided by the particular screening assay, in particular further structural analysis by e.g., ¹⁵N NMR relaxation rate determinations or x-ray crystallography with the modulator bound to the subject polypeptide. These studies may be performed in conjunction with biochemical assays.

Once identified, a potential modulator may be used as a model structure, and analogs to the compound can be obtained. The analogs are then screened for their ability to bind the subject polypeptide. An analog of the potential modulator might be chosen as a modulator when it binds to the subject polypeptide with a higher binding affinity than the predecessor modulator.

In a related approach, iterative drug design is used to identify modulators of a target protein. Iterative drug design is a method for optimizing associations between a protein and a modulator by determining and evaluating the three dimensional structures of successive sets of protein/modulator complexes. In iterative drug design, crystals of a series of protein/modulator complexes are obtained and then the three-dimensional structures of each complex is solved. Such an approach provides insight into the association between the proteins and modulators of each complex. For example, this approach may be accomplished by selecting modulators with inhibitory activity, obtaining crystals of this

new protein/modulator complex, solving the three dimensional structure of the complex, and comparing the associations between the new protein/modulator complex and previously solved protein/modulator complexes. By observing how changes in the modulator affected the protein/modulator associations, these associations may be optimized.

In addition to designing and/or identifying a chemical entity to associate with a druggable region, as described above, the same techniques and methods may be used to design and/or identify chemical entities that either associate, or do not associate, with affinity regions, selectivity regions or undesired regions of protein targets. By such methods, selectivity for one or a few targets, or alternatively for multiple targets, from the same species or from multiple species, can be achieved.

For example, a chemical entity may be designed and/or identified for which the binding energy for one druggable region, e.g., an affinity region or selectivity region, is more favorable than that for another region, e.g., an undesired region, by about 20%, 30%, 50% to about 60% or more. It may be the case that the difference is observed between (a) more than two regions, (b) between different regions (selectivity, affinity or undesirable) from the same target, (c) between regions of different targets, (d) between regions of homologs from different species, or (e) between other combinations. Alternatively, the comparison may be made by reference to the K_d , usually the apparent K_d , of said chemical entity with the two or more regions in question.

In another aspect, prospective modulators are screened for binding to two nearby druggable regions on a target protein. For example, a modulator that binds a first region of a target polypeptide does not bind a second nearby region. Binding to the second region can be determined by monitoring changes in a different set of amide chemical shifts in either the original screen or a second screen conducted in the presence of a modulator (or potential modulator) for the first region. From an analysis of the chemical shift changes, the approximate location of a potential modulator for the second region is identified. Optimization of the second modulator for binding to the region is then carried out by screening structurally related compounds (e.g., analogs as described above). When modulators for the first region and the second region are identified, their location and orientation in the ternary complex can be determined experimentally. On the basis of this structural information, a linked compound, e.g., a consolidated modulator, is synthesized in which the modulator for the first region and the modulator for the second region are linked. In certain embodiments, the two modulators are covalently linked to form a consolidated

modulator. This consolidated modulator may be tested to determine if it has a higher binding affinity for the target than either of the two individual modulators. A consolidated modulator is selected as a modulator when it has a higher binding affinity for the target than either of the two modulators. Larger consolidated modulators can be constructed in an analogous manner, e.g., linking three modulators which bind to three nearby regions on the target to form a multilinked consolidated modulator that has an even higher affinity for the target than the linked modulator. In this example, it is assumed that is desirable to have the modulator bind to all the druggable regions. However, it may be the case that binding to certain of the druggable regions is not desirable, so that the same techniques may be used to identify modulators and consolidated modulators that show increased specificity based on binding to at least one but not all druggable regions of a target.

The present invention provides a number of methods that use drug design as described above. For example, in one aspect, the present invention contemplates a method for designing a candidate compound for screening for inhibitors of a polypeptide of the invention, the method comprising: (a) determining the three dimensional structure of a crystallized polypeptide of the invention or a fragment thereof; and (b) designing a candidate inhibitor based on the three dimensional structure of the crystallized polypeptide or fragment.

In another aspect, the present invention contemplates a method for identifying a potential inhibitor of a polypeptide of the invention, the method comprising: (a) providing the three-dimensional coordinates of a polypeptide of the invention or a fragment thereof; (b) identifying a druggable region of the polypeptide or fragment; and (c) selecting from a database at least one compound that comprises three dimensional coordinates which indicate that the compound may bind the druggable region; (d) wherein the selected compound is a potential inhibitor of a polypeptide of the invention.

In another aspect, the present invention contemplates a method for identifying a potential modulator of a molecule comprising a druggable region similar to that of SEQ ID NO: 2 or SEQ ID NO: 4, the method comprising: (a) using the atomic coordinates of amino acid residues from SEQ ID NO: 2 or SEQ ID NO: 4, or a fragment thereof, \pm a root mean square deviation from the backbone atoms of the amino acids of not more than 1.5 Å, to generate a three-dimensional structure of a molecule comprising a druggable region that is a portion of SEQ ID NO: 2 or SEQ ID NO: 4; (b) employing the three dimensional structure to design or select the potential modulator; (c) synthesizing the modulator; and

(d) contacting the modulator with the molecule to determine the ability of the modulator to interact with the molecule.

In another aspect, the present invention contemplates an apparatus for determining whether a compound is a potential inhibitor of a polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, the apparatus comprising: (a) a memory that comprises: (i) the three dimensional coordinates and identities of the atoms of a polypeptide of the invention or a fragment thereof that form a druggable site; and (ii) executable instructions; and (b) a processor that is capable of executing instructions to: (i) receive three-dimensional structural information for a candidate compound; (ii) determine if the three-dimensional structure of the candidate compound is complementary to the structure of the interior of the druggable site; and (iii) output the results of the determination.

In another aspect, the present invention contemplates a method for designing a potential compound for the prevention or treatment of *P. aeruginosa* related disease or disorder, the method comprising: (a) providing the three dimensional structure of a crystallized polypeptide of the invention, or a fragment thereof; (b) synthesizing a potential compound for the prevention or treatment of *P. aeruginosa* related disease or disorder based on the three dimensional structure of the crystallized polypeptide or fragment; (c) contacting a polypeptide of the present invention or an *P. aeruginosa* with the potential compound; and (d) assaying the activity of a polypeptide of the present invention, wherein a change in the activity of the polypeptide indicates that the compound may be useful for prevention or treatment of a *P. aeruginosa* related disease or disorder.

In another aspect, the present invention contemplates a method for designing a potential compound for the prevention or treatment of *P. aeruginosa* related disease or disorder, the method comprising: (a) providing structural information of a druggable region derived from NMR spectroscopy of a polypeptide of the invention, or a fragment thereof; (b) synthesizing a potential compound for the prevention or treatment of *P. aeruginosa* related disease or disorder based on the structural information; (c) contacting a polypeptide of the present invention or an *P. aeruginosa* with the potential compound; and (d) assaying the activity of a polypeptide of the present invention, wherein a change in the activity of the polypeptide indicates that the compound may be useful for prevention or treatment of a *P. aeruginosa* related disease or disorder.

(b) *In Vitro Assays*

Polypeptides of the invention may be used to assess the activity of small molecules and other modulators in *in vitro* assays. In one embodiment of such an assay, agents are identified which modulate the biological activity of a protein, protein-protein interaction of interest or protein complex, such as an enzymatic activity, binding to other cellular components, cellular compartmentalization, signal transduction, and the like. In certain embodiments, the test agent is a small organic molecule.

Assays may employ kinetic or thermodynamic methodology using a wide variety of techniques including, but not limited to, microcalorimetry, circular dichroism, capillary zone electrophoresis, nuclear magnetic resonance spectroscopy, fluorescence spectroscopy, and combinations thereof.

The invention also provides a method of screening compounds to identify those which modulate the action of polypeptides of the invention, or polynucleotides encoding the same. The method of screening may involve high-throughput techniques. For example, to screen for modulators, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, comprising a polypeptide of the invention and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate molecule that may be a modulator of a polypeptide of the invention. The ability of the candidate molecule to modulate a polypeptide of the invention is reflected in decreased binding of the labeled ligand or decreased production of product from such substrate. Detection of the rate or level of production of product from substrate may be enhanced by using a reporter system. Reporter systems that may be useful in this regard include but are not limited to colorimetric labeled substrate converted into product, a reporter gene that is responsive to changes in a nucleic acid of the invention or polypeptide activity, and binding assays known in the art.

Another example of an assay for a modulator of a polypeptide of the invention is a competitive assay that combines a polypeptide of the invention and a potential modulator with molecules that bind to a polypeptide of the invention, recombinant molecules that bind to a polypeptide of the invention, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. Polypeptides of the invention can be labeled, such as by radioactivity or a colorimetric compound, such that

the number of molecules of a polypeptide of the invention bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential modulator.

A number of methods for identifying a molecule which modulates the activity of a polypeptide are known in the art. For example, in one such method, a subject polypeptide is contacted with a test compound, and the activity of the subject polypeptide in the presence of the test compound is determined, wherein a change in the activity of the subject polypeptide is indicative that the test compound modulates the activity of the subject polypeptide. In certain instances, the test compound agonizes the activity of the subject polypeptide, and in other instances, the test compound antagonizes the activity of the subject polypeptide.

In another example, a compound which modulates SEQ ID NO: 2 or SEQ ID NO: 4 dependent growth or infectivity of *P. aeruginosa* may be identified by (a) contacting a polypeptide of the invention with a test compound; and (b) determining the activity of the polypeptide in the presence of the test compound, wherein a change in the activity of the polypeptide is indicative that the test compound may modulate the growth or infectivity of *P. aeruginosa*.

(c) In Vivo Assays

Animal models of bacterial infection and/or disease may be used as an *in vivo* assay for evaluating the effectiveness of a potential drug target in treating or preventing *P. aeruginosa* related diseases or disorders. A number of suitable animal models are described briefly below, however, these models are only examples and modifications, or completely different animal models, may be used in accord with the methods of the invention.

(i) Mouse Soft Tissue Model

The mouse soft tissue infection model is a sensitive and effective method for measurement of bacterial proliferation. In these models (Vogelman et al., 1988, J. Infect. Dis. 157: 287-298) anesthetized mice are infected with the bacteria in the muscle of the hind thigh. The mice can be either chemically immune compromised (e.g., cytoxan treated at 125 mg/kg on days -4, -2, and 0) or immunocompetent. The dose of microbe necessary to cause an infection is variable and depends on the individual microbe, but commonly is on the order of 10^5 - 10^6 colony forming units per injection for bacteria. A variety of mouse strains are useful in this model although Swiss Webster and DBA2 lines are most

commonly used. Once infected the animals are conscious and show no overt ill effects of the infections for approximately 12 hours. After that time virulent strains cause swelling of the thigh muscle, and the animals can become bacteremic within approximately 24 hours. This model most effectively measures proliferation of the microbe, and this proliferation is measured by sacrifice of the infected animal and counting colonies from homogenized thighs.

(ii) Diffusion Chamber Model

A second model useful for assessing the virulence of microbes is the diffusion chamber model (Malouin et al., 1990, *Infect. Immun.* 58: 1247-1253; Doy et al., 1980, *J. Infect. Dis.* 2: 39-51; Kelly et al., 1989, *Infect. Immun.* 57: 344-350. In this model rodents have a diffusion chamber surgically placed in the peritoneal cavity. The chamber consists of a polypropylene cylinder with semipermeable membranes covering the chamber ends. Diffusion of peritoneal fluid into and out of the chamber provides nutrients for the microbes. The progression of the "infection" may be followed by examining growth, the exoproduct production or RNA messages. The time experiments are done by sampling multiple chambers.

(iii) Endocarditis Model

For bacteria, an important animal model effective in assessing pathogenicity and virulence is the endocarditis model (J. Santoro and M. E. Levinson, 1978, *Infect. Immun.* 19: 915-918). A rat endocarditis model can be used to assess colonization, virulence and proliferation.

(iv) Osteomyelitis Model

A fourth model useful in the evaluation of pathogenesis is the osteomyelitis model (Spagnolo et al., 1993, *Infect. Immun.* 61: 5225-5230). Rabbits are used for these experiments. Anesthetized animals have a small segment of the tibia removed and microorganisms are microinjected into the wound. The excised bone segment is replaced and the progression of the disease is monitored. Clinical signs, particularly inflammation and swelling are monitored. Termination of the experiment allows histologic and pathologic examination of the infection site to complement the assessment procedure.

(v) Murine Septic Arthritis Model

A fifth model relevant to the study of microbial pathogenesis is a murine septic arthritis model (Abdelnour et al., 1993, *Infect. Immun.* 61: 3879-3885). In this model mice are infected intravenously and pathogenic organisms are found to cause inflammation in distal limb joints. Monitoring of the inflammation and comparison of inflammation vs. inocula allows assessment of the virulence of related strains.

(vi) *Bacterial Peritonitis Model*

Finally, bacterial peritonitis offers rapid and predictive data on the virulence of strains (M. G. Bergeron, 1978, *Scand. J. Infect. Dis. Suppl.* 14: 189-206; S. D. Davis, 1975, *Antimicrob. Agents Chemother.* 8: 50-53). Peritonitis in rodents, such as mice, can provide essential data on the importance of targets. The end point may be lethality or clinical signs can be monitored. Variation in infection dose in comparison to outcome allows evaluation of the virulence of individual strains.

A variety of other *in vivo* models are available and may be used when appropriate for specific pathogens or specific test agents. For example, target organ recovery assays (Gordee et al., 1984, *J. Antibiotics* 37:1054-1065; Bannatyne et al., 1992, *Infect.* 20:168-170) may be useful for fungi and for bacterial pathogens which are not acutely virulent to animals.

It is also relevant to note that the species of animal used for an infection model, and the specific genetic make-up of that animal, may contribute to the effective evaluation of the effects of a particular test agent. For example, immuno-incompetent animals may, in some instances, be preferable to immuno-competent animals. For example, the action of a competent immune system may, to some degree, mask the effects of the test agent as compared to a similar infection in an immuno-incompetent animal. In addition, many opportunistic infections, in fact, occur in immuno-compromised patients, so modeling an infection in a similar immunological environment is appropriate.

10. Vaccines

There are provided by the invention, products, compositions and methods for raising immunological response against a pathogen, especially *P. aeruginosa*. In one aspect, a polypeptide of the invention or a nucleic acid of the invention, or an antigenic fragment

thereof, may be administered to a subject, optionally with a booster, adjuvant, or other composition that stimulates immune responses.

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal which comprises inoculating the individual with a polypeptide of the invention and/or a nucleic acid of the invention, adequate to produce antibody and/or T cell immune response to protect said individual from infection, particularly bacterial infection and most particularly *P. aeruginosa* infection. Also provided are methods whereby such immunological response slows bacterial replication. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises delivering to such individual a nucleic acid vector, sequence or ribozyme to direct expression of a polypeptide of the invention and/or a nucleic acid of the invention *in vivo* in order to induce an immunological response, such as, to produce antibody and/or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said individual, preferably a human, from disease, whether that disease is already established within the individual or not. One example of administering the gene is by accelerating it into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may comprise DNA, RNA, a ribozyme, a modified nucleic acid, a DNA/RNA hybrid, a DNA-protein complex or an RNA-protein complex.

A further aspect of the invention relates to an immunological composition that when introduced into an individual, preferably a human, capable of having induced within it an immunological response, induces an immunological response in such individual to a nucleic acid of the invention and/or a polypeptide encoded therefrom, wherein the composition comprises a recombinant nucleic acid of the invention and/or polypeptide encoded therefrom and/or comprises DNA and/or RNA which encodes and expresses an antigen of said nucleic acid of the invention, polypeptide encoded therefrom, or other polypeptide of the invention. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity and/or cellular immunity, such as cellular immunity arising from CTL or CD4+T cells.

In another embodiment, the invention relates to compositions comprising a polypeptide of the invention and an adjuvant. The adjuvant can be any vehicle which would typically enhance the antigenicity of a polypeptide, e.g., minerals (for instance, alum, aluminum hydroxide or aluminum phosphate), saponins complexed to membrane protein

antigens (immune stimulating complexes), pluronic polymers with mineral oil, killed mycobacteria in mineral oil, Freund's complete adjuvant, bacterial products, such as muramyl dipeptide (MDP) and lipopolysaccharide (LPS), as well as lipid A, liposomes, or any of the other adjuvants known in the art. A polypeptide of the invention can be emulsified with, absorbed onto, or coupled with the adjuvant.

A polypeptide of the invention may be fused with co-protein or chemical moiety which may or may not by itself produce antibodies, but which is capable of stabilizing the first protein and producing a fused or modified protein which will have antigenic and/or immunogenic properties, and preferably protective properties. Thus fused recombinant protein, may further comprise an antigenic co-protein, such as lipoprotein D from *Hemophilus influenzae*, Glutathione-S-transferase (GST) or beta-galactosidase, or any other relatively large co-protein which solubilizes the protein and facilitates production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system of the organism receiving the protein. The co-protein may be attached to either the amino- or carboxy-terminus of a polypeptide of the invention.

Provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides and/or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. et al. Science 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof, which have been shown to encode non-variable regions of bacterial cell surface proteins, in polynucleotide constructs used in such genetic immunization experiments in animal models of infection with *P. aeruginosa*. Such experiments will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value, derived from the requisite organ of the animal successfully resisting or clearing infection, for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *P. aeruginosa* infection, in mammals, particularly humans.

A polypeptide of the invention may be used as an antigen for vaccination of a host to produce specific antibodies which protect against invasion of bacteria, for example by blocking adherence of bacteria to damaged tissue.

11. Array Analysis

In part, the present invention is directed to the use of subject nucleic acids in arrays to assess gene expression. In another part, the present invention is directed to the use of subject nucleic acids in arrays for *P. aeruginosa*. In yet another part, the present invention contemplates using the subject nucleic acids to interact with probes contained on arrays.

In one aspect, the present invention contemplates an array comprising a substrate having a plurality of addresses, wherein at least one of the addresses has disposed thereon a capture probe that can specifically bind to a nucleic acid of the invention. In another aspect, the present invention contemplates a method for detecting expression of a nucleotide sequence which encodes a polypeptide of the invention, or a fragment thereof, using the foregoing array by: (a) providing a sample comprising at least one mRNA molecule; (b) exposing the sample to the array under conditions which promote hybridization between the capture probe disposed on the array and a nucleic acid complementary thereto; and (c) detecting hybridization between an mRNA molecule of the sample and the capture probe disposed on the array, thereby detecting expression of a sequence which encodes for a polypeptide of the invention, or a fragment thereof.

Arrays are often divided into microarrays and macroarrays, where microarrays have a much higher density of individual probe species per area. Microarrays may have as many as 1000 or more different probes in a 1 cm² area. There is no concrete cut-off to demarcate the difference between micro- and macroarrays, and both types of arrays are contemplated for use with the invention.

Microarrays are known in the art and generally consist of a surface to which probes that correspond in sequence to gene products (e.g., cDNAs, mRNAs, oligonucleotides) are bound at known positions. In one embodiment, the microarray is an array (e.g., a matrix) in which each position represents a discrete binding site for a product encoded by a gene (e.g., a protein or RNA), and in which binding sites are present for products of most or almost all of the genes in the organism's genome. In certain embodiments, the binding site or site is a nucleic acid or nucleic acid analogue to which a particular cognate cDNA can specifically

hybridize. The nucleic acid or analogue of the binding site may be, e.g., a synthetic oligomer, a full-length cDNA, a less-than full length cDNA, or a gene fragment.

Although in certain embodiments the microarray contains binding sites for products of all or almost all genes in the target organism's genome, such comprehensiveness is not necessarily required. Usually the microarray will have binding sites corresponding to at least 100, 500, 1000, 4000 genes or more. In certain embodiments, arrays will have anywhere from about 50, 60, 70, 80, 90, or even more than 95% of the genes of a particular organism represented. The microarray typically has binding sites for genes relevant to testing and confirming a biological network model of interest. Several exemplary human microarrays are publicly available.

The probes to be affixed to the arrays are typically polynucleotides. These DNAs can be obtained by, e.g., polymerase chain reaction (PCR) amplification of gene segments from genomic DNA, cDNA (e.g., by RT-PCR), or cloned sequences. PCR primers are chosen, based on the known sequence of the genes or cDNA, that result in amplification of unique fragments (e.g., fragments that do not share more than 10 bases of contiguous identical sequence with any other fragment on the microarray). Computer programs are useful in the design of primers with the required specificity and optimal amplification properties. See, e.g., Oligo pl version 5.0 (National Biosciences). In an alternative embodiment, the binding (hybridization) sites are made from plasmid or phage clones of genes, cDNAs (e.g., expressed sequence tags), or inserts therefrom (Nguyen et al., 1995, *Genomics* 29:207-209).

A number of methods are known in the art for affixing the nucleic acids or analogues to a solid support that makes up the array (Schena et al., 1995, *Science* 270:467-470; DeRisi et al., 1996, *Nature Genetics* 14:457-460; Shalon et al., 1996, *Genome Res.* 6:639-645; and Schena et al., 1995, *Proc. Natl. Acad. Sci. USA* 93:10539-11286).

Another method for making microarrays is by making high-density oligonucleotide arrays (Fodor et al., 1991, *Science* 251:767-773; Pease et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:5022-5026; Lockhart et al., 1996, *Nature Biotech* 14:1675; U.S. Pat. Nos. 5,578,832; 5,556,752; and 5,510,270; Blanchard et al., 1996, 11: 687-90).

Other methods for making microarrays, e.g., by masking (Maskos and Southern, 1992, *Nuc. Acids Res.* 20:1679-1684), may also be used. In principal, any type of array, for example, dot blots on a nylon hybridization membrane (see Sambrook et al., *Molecular*

Cloning - A Laboratory Manual (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989), could be used, although, as will be recognized by those of skill in the art.

The nucleic acids to be contacted with the microarray may be prepared in a variety of ways, and may include nucleotides of the subject invention. Such nucleic acids are often labeled fluorescently. Nucleic acid hybridization and wash conditions are chosen so that the population of labeled nucleic acids will specifically hybridize to appropriate, complementary nucleic acids affixed to the matrix. Non-specific binding of the labeled nucleic acids to the array can be decreased by treating the array with a large quantity of non-specific DNA — a so-called “blocking” step.

When fluorescently labeled probes are used, the fluorescence emissions at each site of a transcript array may be detected by scanning confocal laser microscopy. When two fluorophores are used, a separate scan, using the appropriate excitation line, is carried out for each of the two fluorophores used. Fluorescent microarray scanners are commercially available from Affymetrix, Packard BioChip Technologies, BioRobotics and many other suppliers. Signals are recorded, quantitated and analyzed using a variety of computer software.

According to the method of the invention, the relative abundance of an mRNA in two cells or cell lines is scored as a perturbation and its magnitude determined (i.e., the abundance is different in the two sources of mRNA tested), or as not perturbed (i.e., the relative abundance is the same). As used herein, a difference between the two sources of RNA of at least a factor of about 25% (RNA from one source is 25% more abundant in one source than the other source), more usually about 50%, even more often by a factor of about 2 (twice as abundant), 3 (three times as abundant) or 5 (five times as abundant) is scored as a perturbation. Present detection methods allow reliable detection of difference of an order of about 2-fold to about 5-fold, but more sensitive methods are expected to be developed.

In addition to identifying a perturbation as positive or negative, it is advantageous to determine the magnitude of the perturbation. This can be carried out, as noted above, by calculating the ratio of the emission of the two fluorophores used for differential labeling, or by analogous methods that will be readily apparent to those of skill in the art.

In certain embodiments, the data obtained from such experiments reflects the relative expression of each gene represented in the microarray. Expression levels in

different samples and conditions may now be compared using a variety of statistical methods.

12. Pharmaceutical Compositions

Pharmaceutical compositions of this invention include any modulator identified according to the present invention, or a pharmaceutically acceptable salt thereof, and a pharmaceutically acceptable carrier, adjuvant, or vehicle. The term "pharmaceutically acceptable carrier" refers to a carrier(s) that is "acceptable" in the sense of being compatible with the other ingredients of a composition and not deleterious to the recipient thereof.

Methods of making and using such pharmaceutical compositions are also included in the invention. The pharmaceutical compositions of the invention can be administered orally, parenterally, by inhalation spray, topically, rectally, nasally, buccally, vaginally, or via an implanted reservoir. The term parenteral as used herein includes subcutaneous, intracutaneous, intravenous, intramuscular, intra articular, intrasynovial, intrasternal, intrathecal, intralesional, and intracranial injection or infusion techniques.

Dosage levels of between about 0.01 and about 100 mg/kg body weight per day, preferably between about 0.5 and about 75 mg/kg body weight per day of the modulators described herein are useful for the prevention and treatment of disease and conditions, including *P. aeruginosa* mediated diseases and conditions. The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode of administration. A typical preparation will contain from about 5% to about 95% active compound (w/w). Alternatively, such preparations contain from about 20% to about 80% active compound.

13. Antimicrobial Agents

The polypeptides of the invention may be used to develop antimicrobial agents for use in a wide variety of applications. The uses are as varied as surface disinfectants, topical pharmaceuticals, personal hygiene applications (e.g., antimicrobial soap, deodorant or the like), additives to cell culture medium, and systemic pharmaceutical products. Antimicrobial agents of the invention may be incorporated into a wide variety of products

and used to treat an already existing microbial infection/contamination or may be used prophylactically to suppress future infection/contamination.

The antimicrobial agents of the invention may be administered to a site, or potential site, of infection/contamination in either a liquid or solid form. Alternatively, the agent may be applied as a coating to a surface of an object where microbial growth is undesirable using nonspecific absorption or covalent attachment. For example, implants or devices (such as linens, cloth, plastics, heart pacemakers, surgical stents, catheters, gastric tubes, endotracheal tubes, prosthetic devices) can be coated with the antimicrobials to minimize adherence or persistence of bacteria during storage and use. The antimicrobials may also be incorporated into such devices to provide slow release of the agent locally for several weeks during healing. The antimicrobial agents may also be used in association with devices such as ventilators, water reservoirs, air-conditioning units, filters, paints, or other substances. Antimicrobials of the invention may also be given orally or systemically after transplantation, bone replacement, during dental procedures, or during implantation to prevent colonization with bacteria.

In another embodiment, antimicrobial agents of the invention may be used as a food preservative or in treating food products to eliminate potential pathogens. The latter use might be targeted to the fish and poultry industries that have serious problems with enteric pathogens which cause severe human disease. In a further embodiment, the agents of the invention may be used as antimicrobials for food crops, either as agents to reduce post harvest spoilage or to enhance host resistance. The antimicrobials may also be used as preservatives in processed foods either alone or in combination with antibacterial food additives such as lysozymes.

In another embodiment, the antimicrobials of the invention may be used as an additive to culture medium to prevent or eliminate infection of cultured cells with a pathogen.

EXEMPLIFICATION

The invention now being generally described, it will be more readily understood by reference to the following examples which are included merely for purposes of illustration of certain aspects and embodiments of the present invention, and are not intended to limit the invention in any way.

EXAMPLE 1 Isolation and Cloning of Nucleic Acid

Pseudomonas aeruginosa is an opportunistic Gram-negative bacilli found in sewage, plants, and sometimes the intestine. It is capable of infecting various organs and has been identified in numerous infections including those in the ears, lungs, urinary tract, blood and in burns and surgical wound infections. Polynucleotide sequences were obtained from The Institute of Genomic Research (TIGR) (Rockville, MD; www.tigr.org). Chromosomal DNA was acquired from the American Type Culture Collection (ATCC; reference #17933D).

The coding sequence of the polynucleotide having SEQ ID NO: 1 is obtained by reference to either publicly available databases or from the use of a bioinformatics program that is used to select the coding sequence of interest from the genome. For example, coding sequences for the genome of *P. aeruginosa* may be obtained from NCBI (<http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/fragik?db=Genome&gi=163>).

The coding DNA is amplified from purified genomic DNA using PCR with primers that are identified with a computer program. The PCR primers are selected so as to introduce restriction enzyme cleavage sites at the flanking regions of the DNA (e.g., NdeI and BglII). The forward and reverse primers have SEQ ID NO: 5 and SEQ ID NO: 6. The sequences of the primers are shown in FIGURE 5, and their respective restriction sites and melting temperatures are shown in Table 1 of FIGURE 6.

The PCR reaction is performed using 50-100 ng of chromosomal DNA and 2 Units of a high fidelity DNA Polymerase (for example *Pfu* Turbo (Stratagene) or Platinum *Pfx* (Invitrogen)). The thermocycling conditions for the PCR process include a DNA melting step at 94°C for 45 sec, a primer annealing step at 48°C - 58°C (depending on Primer [T_m]) for 45 sec, and an extension step at 68°C - 72°C (depending on enzyme) for 1 min 45 sec - 2 min 30 sec (depending on size of DNA). After 25-30 cycles, a final blocking step at 72°C for 9 min is carried out.

The amplified nucleic acid product is isolated from the PCR cocktail using silica-gel membrane based column chromatography (Qiagen). The quality of the PCR product is assessed by resolving an aliquot of amplified product on a 1% agarose gel. The DNA is quantified spectrophotometrically at A₂₆₀ or by visualizing the resolved genes with a 302 nm UV-B light source.

The PCR product is directionally cloned into the polylinker region of any of three expression vectors: pET28 (Novagen), pET15 (Novagen) or pGEX (Pharmacia/LKB Biotechnology). Additional restriction enzyme sites may be engineered into the expressions vectors to allow for simultaneous clones to be prepared having different purification tags. After the ligation reaction, the DNA is transformed into competent *E. coli* cells (Strains XL1-Blue (Stratagene) or DH5 α (Invitrogen)) via heat shock or electroporation as described in Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). The expression vectors contain the bacteriophage T7 promoter for RNA polymerase, and the *E. coli* strain used produces T7 RNA polymerase upon induction with isopropyl- β -D-thiogalactoside (IPTG). The sequence of the cloning site adds a Glutathione S-transferase (GST) tag, or a polyhistidine (6X His) tag, at the N- or C- terminus of the recombinant protein. The cloning site also inserts a cleavage site for the thrombin or Tev (Invitrogen) enzymes between the recombinant protein and the N- or C- terminal GST or polyhistidine tag.

Transformants are selected using the appropriate antibiotic (Ampicillin or Kanamycin) and identified using PCR, or another method, to analyze their DNA. The polynucleotide sequence cloned into the expression construct is then isolated using a modified alkaline lysis method (Birnboim, H.C., and Doly, J. (1979) *Nucl. Acids Res.* 7, 1513-1522.) The sequence of the clone is verified by standard polynucleotide sequencing methods. The published nucleic acid and amino acid sequences are presented in FIGURE 1 and FIGURE 2. The experimentally determined nucleic acid sequence is presented in FIGURE 3, and the amino acid sequence predicted from the sequence of FIGURE 3 is presented in FIGURE 4.

The expression construct is transformed into a bacterial host strain BL21-Gold (DE3) supplemented with a plasmid called pUBS520, which directs expression of tRNA for arginine (agg and aga) and serves to augment the expression of the recombinant protein in the host cell (Gene, vol. 85 (1989) 109-114). The expression construct may also be transformed into BL21-Gold (DE3) without pUBS520, BL21-Gold (DE3) Codon-Plus (RIL) or (RP) (Stratagene) or Roseatta (DE3) (Novagen), the latter two of which contain genes encoding tRNAs. Alternatively, the expression construct may be transformed into BL21 STAR *E. coli* (Invitrogen) cells which has an Rnase deficiency that reduces degradation of recombinant mRNA transcript and therefore increases the protein yield. The

recombinant protein is then assayed for positive overexpression in the host and the presence of the protein in the cytoplasmic (water soluble) region of the cell.

EXAMPLE 2 Expressing Polypeptides of the Invention

(a) Test Expression

Transformed cells are grown in LB medium supplemented with the appropriate antibiotics up to a final concentration of 100 µg/ml. The cultures are shaken at 37°C until they reach an optical density (OD₆₀₀) between 0.6 and 0.7. The cultures are then induced with isopropyl-beta-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM at 15°C for 10 hours, 25°C for 4 hours, or 30°C for 4 hours.

(b) Method One for Determining Protein Solubility Levels

The cells are harvested by centrifugation and subjected to a freeze/thaw cycle. The cells are lysed using detergent, sonication, or incubation with lysozyme. Total and soluble proteins are assayed using a 26-well BioRad Criterion gel running system. The proteins are stained with an appropriate dye (Coomassie, Silver stain, or Sypro-Red) and visualized with the appropriate visualization system. Typically recombinant protein is seen as a prominent band in the lanes of the gel representing the soluble fraction.

(c) Method Two for Determining Protein Solubility Levels

The soluble and insoluble fractions (in the presence of 6M urea) of the cell pellet are bound to the appropriate affinity column. The purified proteins from both fractions are analysed by SDS-PAGE and the levels of protein in the soluble fraction are determined.

The approximate percent solubility of the polypeptide having the sequence of SEQ ID NO: 4 is determined using one of the foregoing methods, and the resulting percent solubility is presented in Table 1 of FIGURE 6.

(d) Native Protein Expression

The expression construct clone encoding the soluble polypeptide having the amino acid sequence of SEQ ID NO: 4 is introduced into an expression host. The resultant cell line is then grown in culture. The method of growth is dependant on whether the protein to be purified is a native protein or a labeled protein. For native and ¹⁵N labeled protein production, a Gold-pUBS520 (as described above), BL21-Gold (DE3) Codon-Plus (RIL) or

(RP), or BL21 STAR *E. Coli* cell line is used. For generating proteins metabolically labeled with selenium, the clone is introduced into a strain called B834 (Novagen). The methods for expressing labeled polypeptides of the invention are described in the Examples that follow.

In one method for expressing an unlabeled polypeptide of the invention, 2L LB cultures or 1L TB cultures are inoculated with a 1% (v/v) starter culture (OD_{600} of 0.8). The cultures are shaken at 37°C and 200 rpm and grown to an OD_{600} of 0.6-0.8 followed by induction with 0.5mM IPTG at 15°C and 200 rpm for at least 10 hours or at 25°C for 4 hours.

The cells are harvested by centrifugation and the pellets are resuspended in 25 ml HEPES buffer (50 mM, pH 7.5), supplemented with 100 μ l of protease inhibitors (PMSF and benzamidine (Sigma)) and flash-frozen in liquid nitrogen.

Alternatively, for an unlabeled polypeptide of the invention, a starter culture is prepared in a 300 mL Tunair flask (Shelton Scientific) by adding 20 mL of medium having 47.6 g/L of Terrific Broth and 1.5% glycerol in dH₂O followed by autoclaving for 30 minutes at 121°C and 15 psi. When the broth cools to room temperature, the medium is supplemented with 6.3 μ M CoCl₂-6H₂O, 33.2 μ M MnSO₄-5H₂O, 5.9 μ M CuCl₂-2H₂O, 8.1 μ M H₃BO₃, 8.3 μ M Na₂MoO₄-2H₂O, 7 μ M ZnSO₄-7H₂O, 108 μ M FeSO₄-7H₂O, 68 μ M CaCl₂-2H₂O, 4.1 μ M AlCl₃-6H₂O, 8.4 μ M NiCl₂-6H₂O, 1 mM MgSO₄, 0.5% v/v of Kao and Michayluk vitamins mix (Sigma; Cat. No. K3129), 25 μ g/mL Carbenicillin, and 50 μ g/mL Kanamycin. The medium is then inoculated with several colonies of the freshly transformed expression construct of interest. The culture is incubated at 37°C and 260 rpm for about 3 hours and then transferred to a 2.5L Tunair Flask containing 1L of the above media. The 1L culture is then incubated at 37°C with shaking at 230-250 rpm on an orbital shaker having a 1 inch orbital diameter. When the culture reaches an OD_{600} of 3-6 it is induced with 0.5 mM IPTG. The induced culture is then incubated at 15°C with shaking at 230-250 rpm or faster for about 6-15 hours. The cells are harvested by centrifugation at 3500 rpm at 4°C for 20 minutes and the cell pellet is resuspended in 15 mL ice cold binding buffer (Hepes 50 mM, pH 7.5) and 100 μ l of protease inhibitors (50 mM PMSF and 100 mM Benzamidine, stock concentration) and flash frozen.

EXAMPLE 3 Expression of Selmet Labeled Polypeptides

The freshly transformed cell, harboring a plasmid with a nucleic acid encoding a polypeptide of the invention, is inoculated into 20 ml of NMM (New Minimal Medium) and shaken at 37°C for 8-9 hours. This culture is then transferred into a 6L Erlenmeyer flask containing 2L of minimum medium (M9). The media is supplemented with all amino acids except methionine. All amino acids are added as a solution except for Tyrosine, Tryptophan and Phenylalanine which are added to the media in powder format. As well the media is supplemented with MgSO_4 (2mM final concentration), $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ (25mg/L final concentration), Glucose (0.4% final concentration), CaCl_2 (0.1mM final concentration) and Seleno-L-Methionine (40mg/L final concentration). When the OD_{600} of the cell culture reaches 0.8-0.9, IPTG (0.4 mM final concentration) is added to the medium for protein induction, and the cell culture is kept shaking at 15°C for 10 hours. The cells are harvested by centrifugation at 3500 rpm at 4°C for 20 minutes and the cell pellet is resuspended in 15 mL cold binding buffer (Hepes 50 mM, pH 7.5) and 100 μl of protease inhibitors (PMSF and Benzamidine) and flash frozen.

Alternatively, a starter culture is prepared in a 300 mL Tunair flask (Shelton Scientific) by adding 50 mL of sterile medium having 10% 10XM9 (37.4 mM NH_4Cl (Sigma; Cat. No. A4514), 44 mM KH_2PO_4 (Bioshop, Ontario, Canada; Cat. No. PPM 302), 96 mM Na_2HPO_4 (Sigma; Cat. No. S2429256), and 96 mM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ (Sigma; Cat. No. S9390) final concentration), 450 μM alanine, 190 μM arginine, 302 μM asparagine, 300 μM aspartic acid, 330 μM cysteine, 272 μM glutamic acid, 274 μM glutamine, 533 μM glycine, 191 μM histidine, 305 μM isoleucine, 305 μM leucine, 220 μM lysine, 242 μM phenylalanine, 348 μM proline, 380 μM serine, 336 μM threonine, 196 μM tryptophan, 220 μM tyrosine, and 342 μM valine, 204 μM Seleno-L-Methionine (Sigma; Cat. No. S3132), 0.5% v/v of Kao and Michayluk vitamins mix (Sigma; Cat. No. K3129), 2 mM MgSO_4 (Sigma; Cat. No. M7774), 90 μM $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ (Sigma; Cat. No. F8633), 0.4% glucose (Sigma; Cat. No. G-5400), 100 μM CaCl_2 (Bioshop, Ontario, Canada; Cat. No. CCL 302), 50 $\mu\text{g/mL}$ Ampicillin, and 50 $\mu\text{g/mL}$ Kanamycin in dH_2O . The medium is then inoculated with several colonies of *E. coli* B834 cells (Novagen) freshly transformed with an expression construct clone encoding the polypeptide of interest. The culture is then incubated at 37°C and 200 rpm until it reaches an OD_{600} of ~1 and is then transferred to a 2.5L Tunair Flask containing 1L of the above media. The 1L culture is incubated at 37°C

with shaking at 200 rpm until the culture reaches an OD₆₀₀ of 0.6-0.8 and is then induced with 0.5 mM IPTG. The induced culture is incubated overnight at 15°C with shaking at 200 rpm. The cells are harvested by centrifugation at 4200 rpm at 4°C for 20 minutes and the cell pellet is resuspended in 15 mL ice cold binding buffer (Hepes 50 mM, pH 7.5) and 100 µl of protease inhibitors (50 mM PMSF and 100 mM Benzamidine, stock concentration) and flash frozen.

Alternatively, the cell harboring a plasmid with a nucleic acid encoding a polypeptide of the invention is inoculated into 10 ml of M9 minimum medium and kept shaking at 37°C for 8-9 hours. This culture is then transferred into a 2L Baffled Flask (Corning) containing 1L minimum medium. The media is supplemented with all amino acids except methionine. All are added as a solution, except for Phenylalanine, Alanine, Valine, Leucine, Isoleucine, Proline, and Tryptophan which are added to the media in powder format. As well the media is supplemented with MgSO₄ (2mM final concentration), FeSO₄·7H₂O (25 mg/L final concentration), Glucose (0.5% final concentration), CaCl₂ (0.1 mM final concentration) and Seleno-Methionine (50 mg/L final concentration). When the OD₆₀₀ of the cell culture reaches 0.8-0.9, IPTG (0.8 mM final concentration) is added to the medium for protein induction, and the cell culture is kept shaking at 25°C for 4 hours. The cells are harvested by centrifuged at 3500 rpm at 4°C for 20 minutes and the cell pellet is resuspended in 10 mL cold binding buffer (Hepes 50 mM, pH 7.5) and 100 µl of protease inhibitors (PMSF and Benzamidine) and flash frozen.

EXAMPLE 4 Expression of ¹⁵N Labeled Polypeptides

The cell, harboring a plasmid with a nucleic acid encoding a polypeptide of the invention, is inoculated into 2L of minimal media (containing ¹⁵N isotope, Cambridge Isotope Lab) in a 6L Erlenmeyer flask. The minimal media is supplemented with 0.01 mM ZnSO₄, 0.1 mM CaCl₂, 1 mM MgSO₄, 5 mg/L Thiamine.HCl, and 0.4% glucose. The 2L culture is grown at 37°C and 200 rpm to an OD₆₀₀ of between 0.7-0.8. The culture is then induced with 0.5 mM IPTG and allowed to shake at 15°C for 14 hours. The cells are harvested by centrifugation and the cell pellet is resuspended in 15 mL cold binding buffer and 100µl of protease inhibitor and flash frozen. The protein is then purified as described below.

Alternatively, the freshly transformed cell, harboring a plasmid with the gene of interest, is inoculated into 10 mL of M9 media (with ^{15}N isotope) and supplemented with 0.01 mM ZnSO_4 , 0.1 mM CaCl_2 , 1 mM MgSO_4 , 5 mg/L Thiamine.HCl, and 0.4% glucose. After 8-10 hours of growth at 37°C, the culture is transferred to a 2L Baffled flask (Corning) containing 990 mL of the same media. When OD_{600} of the culture is between 0.7-0.8, protein production is initiated by adding IPTG to a final concentration of 0.8 mM and lowering the temperature to 25°C. After 4 hours of incubation at this temperature, the cells are harvested, and the cell pellet is resuspended in 10 mL cold binding buffer (Hepes 50 mM, pH 7.5) and 100 μL of protease inhibitor and flash frozen.

EXAMPLE 5 Method One for Purifying Polypeptides of the Invention

The frozen pellets are thawed and sonicated to lyse the cells (5 x 30 seconds, output 4 to 5, 80% duty cycle, in a Branson Sonifier, VWR). The lysates are clarified by centrifugation at 14,000 rpm for 60 min at 4°C to remove insoluble cellular debris. The supernatants are removed and supplemented with 1 μL of Benzonase Nuclease (25 U/ μL , Novagen).

The recombinant protein is purified using DE52 (anion exchanger, Whatman) and Ni-NTA columns (Qiagen). The DE52 columns (30 mm wide, Biorad) are prepared by mixing 10 grams of DE52 resin in 25 ml of 2.5 M NaCl per protein sample, applying the resin to the column and equilibrating with 30 ml of binding buffer (50 mM in HEPES, pH 7.5, 5% glycerol (v/v), 0.5 M NaCl, 5 mM imidazole). Ni-NTA columns are prepared by adding 3.5-8 ml of resin to the column (20 mm wide, Biorad) based on the level of expression of the recombinant protein and equilibrating the column with 30 ml of binding buffer. The columns are arranged in tandem so that the protein sample is first passed over the DE52 column and then loaded directly onto the Ni-NTA column.

The Ni-NTA columns are washed with at least 150 ml of wash buffer (50mM HEPES, pH 7.5, 5% glycerol (v/v), 0.5 M NaCl, 30 mM imidazole) per column. A pump may be used to load and/or wash the columns. The protein is eluted off of the Ni-NTA column using elution buffer (50 mM in HEPES, pH 7.5, 5% glycerol (v/v), 0.5 M NaCl, 250 mM imidazole) until no more protein is observed in the aliquots of eluate as measured using Bradford reagent (Biorad). The eluate is supplemented with 1 mM of EDTA and 0.2 mM DTT.

The samples are assayed by SDS-PAGE and stained with Coomassie Blue, with protein purity determined by visual staining.

Two methods may be used to remove the His tag located at either the C or N-terminus. In certain instances, the His tag may not be removed. In either case, the expressed polypeptide will have additional residues attributable to the His tag, as shown in the following table:

| <i>SEQ ID NO for Additional Residues</i> | <i>Additional Residues</i> | <i>Type of Tag and Whether or Not Removed</i> |
|----------------------------------------------|----------------------------|---------------------------------------------------|
| N/A | GSH | His tag removed from N-terminus |
| SEQ ID NO: 7 | MGSSHHHHHHSSGLVPRG SH | His tag not removed from N-terminus |
| SEQ ID NO: 8 | GSENLVFQGHHHHHH | His tag removed from C-terminus |
| SEQ ID NO: 9 | GSENLVFQ | His tag not removed from C-terminus |

In method one, a sample of purified polypeptide are supplemented with 2.5 mM CaCl_2 and an appropriate amount of thrombin (the amount added will vary depending on the activity of the enzyme preparation) and incubated for ~20-30 minutes on ice in order to remove the His tag. In method two, a sample of purified polypeptide is combined with thirty units of recombinant TEV protease in 50 mmol TRIS HCl pH = 8.0, 0.5 mmol EDTA and 1 mmol DTT, followed by incubation at 4°C overnight, to remove the His tag.

The protein sample is then dialyzed in dialysis buffer (10mM HEPES, pH 7.5, 5% glycerol (v/v) and 0.5 M NaCl) for at least 8 hours using a Slide-A-Lyzer (Pierce) appropriate for the molecular weight of the recombinant protein. An aliquot of the cleaved and dialyzed samples is then assayed by SDS-PAGE and stained with Coomassie Blue to determine the purity of the protein and the success of cleavage.

The remainder of the sample is centrifuged at 2700 rpm at 4°C for 10-15 minutes to remove any precipitant and supplemented with 100 μl of protease inhibitor cocktail (0.1 M benzamidine and 0.05 M PMSF) (NO Bioshop). The protein is then applied to a second Ni-NTA column (~8 ml of resin) to remove the His-tags and eluted with binding buffer or wash buffer until no more protein is eluting off the column as assayed using the Bradford reagent. The eluted sample is supplemented with 1 mM EDTA and 0.6 mM of DTT and concentrated to a final volume of ~15 mls using a Millipore Concentrator with an

appropriately sized filter at 2700 rpm at 4°C. The samples are then dialyzed overnight against crystallization buffer and concentrated to final volume of 0.3-0.7 ml.

EXAMPLE 6 Method Two for Purifying Polypeptides of the Invention

The frozen pellets are thawed and supplemented with 100 µl of protease inhibitor (0.1 M benzamidine and 0.05 M PMSF), 0.5% CHAPS, and 4 U/ml Benzonase Nuclease. The sample is then gently rocked on a Nutator (VWR, setting 3) at room temperature for 30 minutes. The cells are then lysed by sonication (1 x 30 seconds, output 4 to 5, 80% duty cycle, in a Branson Sonifier, VWR) and an aliquot is saved for a gel sample.

The recombinant protein is purified using a three column system. The columns are set up in tandem so that the lysate flows from a Biorad Econo (5.0 x 30 cm x 589 ml) "lysate" column onto a Biorad Econo (2.5 x 20 cm x 98 ml) DE52 column and finally onto a Biorad Econo (1.5 x 15 cm x 27 ml) Ni-NTA column. The lysate is mixed with 10 g of equilibrated DE52 resin and diluted to a total volume of 300 ml with binding buffer. This mixture is poured into the first column which is empty. The remainder of the purification procedure is described in EXAMPLE 5 above.

EXAMPLE 7 Method Three for Purifying Polypeptides of the Invention

The frozen pellets are thawed and sonicated to lyse the cells (5 x 30 seconds, output 4 to 5, 80% duty cycle, in a Branson Sonifier, VWR). The lysates are clarified by centrifugation at 14000 rpm for 60 min at 4°C to remove insoluble cellular debris. The supernatants are removed and supplemented with 1 µl of Benzonase Nuclease (25 U/µl, Novagen).

The recombinant protein is purified using DE52 (anion exchanger, Whatman) and Glutathione sepharose columns (Glutathione-Superflow resin, Clontech). The DE52 columns (30 mm wide, Biorad) are prepared by mixing 10 grams of DE52 resin in 20 ml of 2.5 M NaCl per protein sample, applying the resin to the column and equilibrating with 30 ml of loading buffer (50mM in HEPES, pH 7.5, 10% glycerol (v/v), 0.5 M NaCl, 1 mM EDTA, 1 mM DTT). Glutathione sepharose columns are prepared by adding 3 ml of resin to the column (20 mm wide, Biorad) and equilibrating the column with 30 ml of loading

buffer. The columns are arranged in tandem so that the protein sample is first passed over the DE52 column and then loads directly onto the Glutathione sepharose column.

The columns are washed with at least 150 ml of loading buffer supplemented with protease inhibitor cocktail (0.1 M benzamidine and 0.05 M PMSF) per column. A pump may be used to load and/or wash the columns. The protein is eluted off of the Glutathione sepharose column using elution buffer (20mM HEPES, pH 7.5, 0.5 M NaCl, 1 mM EDTA, 1 mM DTT; 25 mM glutathione (reduced form)) until no more protein is observed in the aliquots of eluate as measured using Biorad Bradford reagent.

The GST tag may be removed using thrombin or other procedures known in the art. The protein samples are then dialyzed into crystallization buffer (10 mM Hepes, pH 7.5, 500 mM NaCl) to remove free glutathione and assayed by SDS-PAGE followed by staining with Coomassie blue. Prior to use or storage, the samples are concentrated to final volume of 0.3-0.5 ml.

Using one or more of the methods described above, purified polypeptide having SEQ ID NO: 4 is obtained in a yield of approximately 49.3 per liter of culture. The purified polypeptide is essentially the only protein visualized in the SDS-PAGE assay using Coomassie Blue described above, which is at least about 95% or greater purity. The polypeptide so expressed and purified is His tagged (having sequence MGSSHHHHHHSSGLVPRGSH as described above) at the N-terminus.

Using one or more of the methods described above, purified selmet labeled polypeptide having SEQ ID NO: 4 is obtained in a yield of approximately 45.8 mg per liter of culture. The purified polypeptide is essentially the only protein visualized in the SDS-PAGE assay using Coomassie Blue described above, which is at least about 95% or greater purity. The polypeptide so expressed and purified is His tagged (having sequence MGSSHHHHHHSSGLVPRGSH) at the N-terminus.

The protein samples so prepared and purified may be used in the biophysical studies that follow, with or without the His tag or the residual amino acids resulting from removal of the His tag. In certain instances, such as EXAMPLE 10, the polypeptide used may be a fusion protein with a specific tag.

A stable solution of purified polypeptide having SEQ ID NO: 4, prepared and purified as described above, may be prepared with 123.2 mg (or a lesser amount) of protein

in one ml of either the dialysis or crystallization buffers (or possibly both) described above in EXAMPLE 5 or EXAMPLE 7, respectively.

A stable solution of purified selmet labeled polypeptide having SEQ ID NO: 4, prepared and purified as described above, may be prepared with 91.6 mg (or a lesser amount) of protein in one ml of either the dialysis or crystallization buffers (or possibly both) described above in EXAMPLE 5 or EXAMPLE 7, respectively.

Certain of the foregoing information is also set forth in Table 1 of FIGURE 6.

EXAMPLE 8 Mass Spectrometry Analysis via Fingerprint Mapping

A gel slice from a purification protocol described above containing a polypeptide of the invention is cut into 1 mm cubes and 10 to 20 μ l of 1% acetic acid is added. After washing with 100 - 150 μ l HPLC grade water and removal of the liquid, acetonitrile (~200 μ l, approximately 3 to 4 times the volume of the gel particles) is added followed by incubation at room temperature for 10 to 15 minutes with vortexing. A second acetonitrile wash may be required to completely dehydrate the gel particles. The protein in the gel particles is reduced at 50 degrees Celsius using 10 mM dithiothreitol (in 100 mM ammonium bicarbonate) and then alkylated at room temperature in the dark using 55 mM iodoacetamide (in 100 mM ammonium bicarbonate). The gel particles are rinsed with a minimal volume of 100 mM ammonium bicarbonate before a trypsin (50 mM ammonium bicarbonate, 5 mM CaCl_2 , and 12.5 ng/ μ l trypsin) solution is added. The gel particles are left on ice for 30 to 45 minutes (after 20 minutes incubation more trypsin solution is added). The excess trypsin solution is removed and 10 to 15 μ l digestion buffer without trypsin is added to ensure the gel particles remain hydrated during digestion. After digestion at 37°C, the supernatant is removed from the gel particles. The peptides are extracted from the gel particles with 2 changes of 100 μ L of 100 mM ammonium bicarbonate with shaking for 45 minutes and pooled with the initial gel supernatant. The extracts are acidified to 1% (v/v) with 100% acetic acid.

The tryptic peptides are purified with a C18 reverse phase resin. 250 μ L of dry resin is washed twice with methanol and twice with 75% acetonitrile/1% acetic acid. A 5:1 slurry of solvent:resin is prepared with 75% acetonitrile/1% acetic acid. To the extracted peptides, 2 μ L of the resin slurry is added and the solution is shaken for 30 minutes at room

temperature. The supernatant is removed and replaced with 200 μ L of 2% acetonitrile/1% acetic acid and shaken for 5-15 minutes. The supernatant is removed and the peptides are eluted from the resin with 15 μ L of 75% acetonitrile/1% acetic acid with shaking for about 5 minutes. The peptide and slurry mixture is applied to a filter plate and centrifuged, and the filtrate is collected and stored at -70°C until use.

Alternatively, the tryptic peptides are purified using ZipTip_{C18} (Millipore, Cat # ZTC18S960). The ZipTips are first pre-wetted by aspirating and dispensing 100% methanol. The tips are then washed with 2% acetonitrile/1% acetic acid (5 times), followed by 65% acetonitrile/1% acetic (5 times) and returned to 2% acetonitrile/1% acetic acid (10 times). The digested peptides are bound to the ZipTips by aspirating and dispensing the samples 5 times. Salts are removed by washing ZipTips with 2% acetonitrile/1% acetic acid (5 times). 10 μ L of 65% acetonitrile/1% acetic acid is collected by the ZipTips and dispensed into a 96-well microtitre plate.

Analytical samples containing tryptic peptides are subjected to MALDI-TOF mass spectrometry. Samples are mixed 1:1 with a matrix of α -cyano-4-hydroxy-*trans*-cinnamic acid. The sample/matrix mixture is spotted on to the MALDI sample plate with a robot, either a Gilson 215 liquid handler or BioMek FX laboratory automation workstation (Beckman). The sample/matrix mixture is allowed to dry on the plate and is then introduced into the mass spectrometer. Analysis of the peptides in the mass spectrometer is conducted using both delayed extraction mode (400 ns delay) and an ion reflector to ensure high resolution of the peptides.

Internally-calibrated tryptic peptide masses are searched against databases using a correlative mass matching algorithm. The Proteometrics software package (ProteoMetrics) is utilized for batch database searching of tryptic peptide mass spectra. Statistical analysis is performed on each protein match to determine its validity. Typical search constraints include error tolerances within 0.1 Da for monoisotopic peptide masses, carboxyamidomethylation of cysteines, no oxidation of methionines allowed, and 0 or 1 missed enzyme cleavages. The software calculates the probability that a candidate in the database search is the protein being analyzed, which is expressed as the Z-score. The Z-score is the distance to the population mean in unit of standard deviation and corresponds to the percentile of the search in the random match population. If a search is in the 95th percentile, for example, about 5% of random matches could yield a higher Z-score than the

search. A Z-score of 1.282 for a search indicates that the search is in the 90th percentile, a Z-score of 1.645 indicates that the search is in the 95th percentile, a Z-score of 2.326 indicates that the search is in the 99th percentile, and a Z-score of 3.090 indicates that the search is in the 99.9th percentile.

EXAMPLE 9 Mass Spectrometry Analysis via High Mass

A matrix solution of 25 mg/mL of 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid) in 66% (v/v) acetonitrile/1% (v/v) acetic acid is prepared along with an internal calibrant of carbonic anhydrase. On to a stainless steel polished MALDI target, 1.5 μ L of a protein solution (concentration of 2 μ g/ μ L) is spotted, followed immediately by 1.5 μ L of matrix. 3 μ L of 40% (v/v) acetonitrile/1% (v/v) acetic acid is then added to each spot has dried. The sample is either spotted manually or utilizing a Gilson 215 liquid handler or BioMek FX laboratory automation workstation (Beckman). The MALDI-TOF instrument utilizes positive ion and linear detection modes. Spectra are acquired automatically over a mass to charge range from 0-150,000 Da, pulsed ion extraction delay is set at 200 ns, and 600 summed shots of 50-shot steps are completed.

The theoretical molecular weight of the protein for MALDI-TOF is determined from its amino acid sequence, taking into account any purification tag or residue thereof still present and any labels (e.g., selenomethionine or ^{15}N). To account for ^{15}N incorporation, an amount equal to the theoretical molecular weight of the protein divided by 70 is added. The mass of water is subtracted from the overall molecular weight. The MALDI-TOF spectrum is calibrated with the internal calibrant of carbonic anhydrase (observed as either $[\text{MH}^+_{\text{avg}}]$ 29025 or $[\text{MH}_2^{2+}]$ 14513).

EXAMPLE 10 Method One for Isolating and Identifying Interacting Proteins

(a) Method One for Preparation of Affinity Column

Micro-columns are prepared using forceps to bend the ends of P200 pipette tips and adding 10 μ L of glass beads to act as a column frit. Six micro-columns are required for every polypeptide to be studied. The micro-columns are placed in a 96-well plate that has 1 mL wells. Next, a series of solutions of the polypeptide having SEQ ID NO: 4 or other polypeptide of the invention, prepared and purified as described above and with a GST tag

on either terminus, is prepared so as to give final amounts of 0, 0.1, 0.5, 1.0, and 2.0 mg of ligand per ml of resin volume.

A slurry of Glutathione-Sepharose 4B (Amersham) is prepared and 0.5 ml slurry/ligand is removed (enough for six 40- μ g aliquots of resin). Using a glass frit Buchner funnel, the resin is washed sequentially with three 10 ml portions each of distilled H₂O and 1 M ACB (20 mM HEPES pH 7.9, 1 M NaCl, 10% glycerol, 1 mM DTT, and 1 mM EDTA). The Glutathione-Sepharose 4B is completely drained of buffer, but not dried. The Glutathione-Sepharose 4B is resuspended as a 50% slurry in 1 M ACB and 80 μ l is added to each micro-column to obtain 40 μ g/column. The buffer containing the ligand concentration series is added to the columns and allowed to flow by gravity. The resin and ligand are allowed to cross-link overnight at 4°C. In the morning, micro-columns are washed with 100 μ l of 1 M ACB and allowed to flow by gravity. This is repeated twice more and the elutions are tested for cross-linking efficiency by measuring the amount of unbound ligand. After washing, the micro-columns are equilibrated using 200 μ l of 0.1 M ACB (20 mM HEPES pH 7.5, 0.1 M NaCl, 10% glycerol, 1 mM DTT, 1 mM EDTA).

In another method, the recombinant GST fusion protein can be replaced by a hexahistidine fusion peptide for use with NTA-Agarose (Qiagen) as the solid support. No adaptation to the above protocol is required for the substitution of NTA agarose for GST Sepharose except that the recombinant protein requires a six histidine fusion peptide in place of the GST fusion.

(b) Method Two for Preparation of Affinity Column

In an alternative method, GST-Sepharose 4B may be replaced by Affi-gel 10 Gel (Bio-Rad). The column resin for affinity chromatography could also be Affigel 10 resin which allows for covalent attachment of the protein ligand to the micro affinity column. An adaptation to the above protocol for the use of this resin is a pre-wash of the resin with 100% isopropanol. No fusion peptides or proteins are required for the use of Affigel 10 resin.

(c) Method One for Bacterial Extract Preparation

A *P. aeruginosa* extract is prepared from cell pellets using a French press followed by sonication. An *P. aeruginosa* cell pellet (~6 g) is suspended in 3 pellet volumes (~20 ml final volume) of 20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 10 mM MgSO₄, 10

mM CaCl_2 , 1 mM DTT, 1 mM PMSF, 1 mM benzamidine, 40 $\mu\text{g/ml}$ RNase A, 75 units/ml S1 nuclease, and 40 units/ml DNase 1. The cell suspension is lysed with one pass with a French Pressure Cell followed by sonication on ice using three bursts of 20 seconds each. The lysate is agitated at 4°C for 30 minutes, brought up to 0.5 M NaCl and then incubated for an additional 30 min at 4°C with agitation. The lysate is centrifuged at 25,000 rpm for 1 hr at 4°C in a Ti70 fixed angle Beckman rotor. The supernatant is removed and dialyzed overnight in a 10,000 Mr dialysis membrane against dialysis buffer (20 mM HEPES pH 7.5, 10 % glycerol, 1 mM DTT, 1 mM EDTA, 100 mM NaCl, 10 mM MgSO_4 , 10 mM CaCl_2 , 1 mM benzamidine, and 1 mM PMSF). The dialyzed protein extract is removed from the dialysis tubing and frozen in one ml aliquots at -70°C.

(d) Method Two for Bacterial Extract Preparation

Bacterial cell extracts from *P. aeruginosa* are prepared from cell pellets using a Bead-Beater apparatus (Bio-spec Products Inc.) and zirconia beads (0.1 mm diameter). The bacterial cell pellet is suspended (~6 g) is suspended in 3 pellet volumes (~20 ml final volume) of 20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 10 mM MgSO_4 , 10 mM CaCl_2 , 1 mM DTT, 1 mM PMSF, 1 mM benzamidine, 40 $\mu\text{g/ml}$ RNase A, 75 units/ml S1 nuclease, and 40 units/ml DNase 1. The cells are lysed with 10 pulses of 30 sec between 90 sec pauses at a temperature of -5 °C. The lysate is separated from the zirconia beads using a standard column apparatus. The lysate is centrifuged at 20000 rpm (48000 x g) in a Beckman JA25.50 rotor. The supernatant is removed and dialyzed overnight at 4 °C against dialysis buffer (20 mM HEPES pH 7.5, 10 % glycerol, 1 mM DTT, 1 mM EDTA, 100 mM NaCl, 10 mM MgSO_4 , 10 mM CaCl_2 , 1 mM benzamidine, and 1 mM PMSF). The dialyzed protein extract is removed from the dialysis tubing and frozen in one ml aliquots at -70°C.

(e) HeLa Cell Extract Preparation

A HeLa cell extract is prepared in the presence of protease inhibitors. Approximately 30 g of Hela cells are submitted to a freeze/thaw cycle and then divided into two tubes. To each tube 20 ml of Buffer A (10 mM HEPES pH 7.9, 1.5 mM MgCl_2 , 10 mM KCl, 0.5 mM DTT, 0.5 mM PMSF) and a protease inhibitor cocktail are added. The cell suspension is homogenized with 10 strokes (2 x 5 strokes) to lyse the cells. Buffer B (15 ml per tube) is added (50 mM HEPES pH 7.9, 1.5 mM MgCl_2 , 1.26 M NaCl, 0.5 mM DTT, 0.5 mM PMSF, 0.5 mM EDTA, 75% glycerol) to each tube followed by a second round of homogenization (2 x 5 strokes). The lysates are stirred on ice for 30 minutes followed by

centrifugation 37,000 rpm for 3 hr at 4°C in a Ti70 fixed angle Beckman rotor. The supernatant is removed and dialyzed overnight in a 10,000 Mr dialysis membrane against dialysis buffer (20 mM HEPES pH 7.9, 10 % glycerol, 1 mM DTT, 1 mM EDTA, and 1 M NaCl. The dialyzed protein extract is removed from the dialysis tubing and frozen in one ml aliquots at -70°C.

(f) Affinity Chromatography

Cell extract is thawed and diluted to 5 mg/ml prior to loading 5 column volumes onto each micro-column. Each column is washed with 5 column volumes of 0.1 M ACB. This washing is repeated once. Each column is then washed with 5 column volumes of 0.1 M ACB containing 0.1% Triton X-100. The columns are eluted with 4 column volumes of 1% sodium dodecyl sulfate into a 96 well PCR plate. To each eluted fraction is added one-tenth volume of 10-fold concentrated loading buffer for SDS-PAGE.

(g) Resolution of the Eluted Proteins and Detection of Bound Proteins

The components of the eluted samples are resolved on SDS-polyacrylamide gels containing 13.8% polyacrylamide using the Laemmli buffer system and stained with silver nitrate. The bands containing the interacting protein are excised with a clean scalpel. The gel volume is kept to a minimum by cutting as close to the band as possible. The gel slice is placed into one well of a low protein binding, 96-well round-bottom plate. To the gel slices is added 20 µl of 1% acetic acid.

EXAMPLE 11 Method Two for Isolating and Identifying Interacting Proteins

Interacting proteins may be isolated using immunoprecipitation. Naturally-occurring bacterial or eukaryotic cells are grown in defined growth conditions or the cells can be genetically manipulated with a protein expression vector. The protein expression vector is used to transiently transfect the cDNA of interest into eukaryotic or prokaryotic cells and the protein is expressed for up to 24 or 48 hours. The cells are harvested and washed three times in sterile 20 mM HEPES (pH7.4)/Hanks balanced salts solution (H/H). The cells are finally resuspended in culture media and incubated at 37°C for 4-8 hr.

The harvested cells may be subjected to one or more culture conditions that may alter the protein profile of the cells for a given period of time. The cells are collected and washed with ice-cold H/H that includes 10 mM sodium pyrophosphate, 10 mM sodium

fluoride, 10mM EDTA, and 1mM sodium orthovanadate. The cells are then lysed in lysis buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% Triton X-100, 10 mM sodium pyrophosphate, 10mM sodium fluoride, 10mM EDTA, 1mM sodium orthovanadate, 1ug/mL PMSF, 1ug/mL aprotinin, 1ug/mL leupetin, and 1ug/mL pepstatin A) by gentle mixing, and placed on ice for 5 minutes. After lysis, the lysate is transferred to centrifuge tubes and centrifuged in an ultracentrifuge at 75000 rpm for 15 min at 4°C. The supernatant is transferred to eppendorf tubes and pre-cleared with 10 µl of rabbit pre-immune antibody on a rotator at 4°C for 1 hr. Forty µl of protein A-Sepharose (Amersham) is then added and incubated at 4°C overnight on a rotator.

The protein A-Sepharose beads are harvested and the supernatant removed to a fresh eppendorf tube. Immune antibody is added to supernatant and rotated for 1 hr at 4°C. Thirty µl of protein A-Sepharose is then added and the mixture is further rotated at 4°C for 1 hr. The beads are harvested and the supernatant is aspirated. The beads are washed three times with 50 mM Tris (pH 8.0), 150 mM NaCl, 0.1% Triton X-100, 10 mM sodium fluoride, 10 mM sodium pyrophosphate, 10 mM sodium orthovanadate, and 10 mM EDTA. Dry the beads with a 50 µl Hamilton syringe. Laemmli loading buffer containing 100 mM DTT is added to the beads and samples are boiled for 5 min. The beads are spun down and the supernatant is loaded onto SDS-PAGE gels. Comparison of the control and experimental samples allows for the selection of polypeptides that interact with the protein of interest.

EXAMPLE 12 Sample for Mass Spectrometry of Interacting Proteins

The gel slices are cut into 1 mm cubes and 10 to 20 µl of 1% acetic acid is added. The gel particles are washed with 100 - 150 µl of HPLC grade water (5 minutes with occasional mixing), briefly centrifuged, and the liquid is removed. Acetonitrile (~200 µl, approximately 3 to 4 times the volume of the gel particles) is added followed by incubation at room temperature for 10 to 15 minutes with vortexing. A second acetonitrile wash may be required to completely dehydrate the gel particles. The sample is briefly centrifuged and all the liquid is removed.

The protein in the gel particles is reduced at 50 degrees Celsius using 10 mM dithiothreitol (in 100 mM ammonium bicarbonate) for 30 minutes and then alkylated at

room temperature in the dark using 55 mM iodoacetamide (in 100 mM ammonium bicarbonate). The gel particles are rinsed with a minimal volume of 100 mM ammonium bicarbonate before a trypsin (50 mM ammonium bicarbonate, 5 mM CaCl_2 , and 12.5 ng/ μL trypsin) solution is added. The gel particles are left on ice for 30 to 45 minutes (after 20 minutes incubation more trypsin solution is added). The excess trypsin solution is removed and 10 to 15 μL digestion buffer without trypsin is added to ensure the gel particles remain hydrated during digestion. The samples are digested overnight at 37°C.

The following day, the supernatant is removed from the gel particles. The peptides are extracted from the gel particles with 2 changes of 100 μL of 100 mM ammonium bicarbonate with shaking for 45 minutes and pooled with the initial gel supernatant. The extracts are acidified to 1% (v/v) with 100% acetic acid.

(a) Method One for Purification of Tryptic Peptides

The tryptic peptides are purified with a C18 reverse phase resin. 250 μL of dry resin is washed twice with methanol and twice with 75% acetonitrile/1% acetic acid. A 5:1 slurry of solvent : resin is prepared with 75% acetonitrile/1% acetic acid. To the extracted peptides, 2 μL of the resin slurry is added and the solution is shaken at moderate speed for 30 minutes at room temperature. The supernatant is removed and replaced with 200 μL of 2% acetonitrile/1% acetic acid and shaken for 5-15 minutes with moderate speed. The supernatant is removed and the peptides are eluted from the resin with 15 μL of 75% acetonitrile/1% acetic acid with shaking for about 5 minutes. The peptide and slurry mixture is applied to a filter plate and centrifuged for 1-2 minutes at 1000 rpm, the filtrate is collected and stored at -70°C until use.

(b) Method Two for Purification of Tryptic Peptides

Alternatively, the tryptic peptides may be purified using ZipTip_{C18} (Millipore, Cat # ZTC18S960). The ZipTips are first pre-wetted by aspirating and dispensing 100% methanol 5 times. The tips are then washed with 2% acetonitrile/1% acetic acid (5 times), followed by 65% acetonitrile/1% acetic (5 times) and returned to 2% acetonitrile/1% acetic acid (5 times). The ZipTips are replaced in their rack and the residual solvent is eliminated. The ZipTips are washed again with 2% acetonitrile/1% acetic acid (5 times). The digested peptides are bound to the ZipTips by aspirating and dispensing the samples 5 times. Salts are removed by washing ZipTips with 2% acetonitrile/1% acetic acid (5 times). 10 μL of

65% acetonitrile/1% acetic acid is collected by the ZipTips and dispensed into a 96-well microtitre plate. 1 μ L of sample and 1 μ L of matrix are spotted on a MALDI-TOF sample plate for analysis.

EXAMPLE 13 Mass Spectrometric Analysis of Interacting Proteins

(a) Method One for Analysis of Tryptic Peptides

Analytical samples containing tryptic peptides are subjected to Matrix Assisted Laser Desorption/Ionization Time Of Flight (MALDI-TOF) mass spectrometry. Samples are mixed 1:1 with a matrix of α -cyano-4-hydroxy-*trans*-cinnamic acid. The sample/matrix mixture is spotted on to the MALDI sample plate with a robot. The sample/matrix mixture is allowed to dry on the plate and is then introduced into the mass spectrometer. Analysis of the peptides in the mass spectrometer is conducted using both delayed extraction mode and an ion reflector to ensure high resolution of the peptides.

Internally-calibrated tryptic peptide masses are searched against both in-house proprietary and public databases using a correlative mass matching algorithm. Statistical analysis is performed on each protein match to determine its validity. Typical search constraints include error tolerances within 0.1 Da for monoisotopic peptide masses and carboxyamidomethylation of cysteines. Identified proteins are stored automatically in a relational database with software links to SDS-PAGE images and ligand sequences.

(b) Method Two for Analysis of Tryptic Peptides

Alternatively, samples containing tryptic peptides are analyzed with an ion trap instrument. The peptide extracts are first dried down to approximately 1 μ L of liquid. To this, 0.1% trifluoroacetic acid (TFA) is added to make a total volume of approximately 5 μ L. Approximately 1-2 μ L of sample are injected onto a capillary column (C8, 150 μ m ID, 15 cm long) and run at a flow rate of 800 nL/min. using the following gradient program:

| Time (minutes) | % Solvent A | % Solvent B |
|----------------|-------------|-------------|
| 0 | 95 | 5 |
| 30 | 65 | 35 |
| 40 | 20 | 80 |
| 41 | 95 | 5 |

Where Solvent A is composed of water/0.5% acetic acid and Solvent B is acetonitrile/0.5% acetic acid. The majority of the peptides will elute between the 20-40 %

acetonitrile gradient. Two types of data from the eluting HPLC peaks are acquired with the ion trap mass spectrometer. In the MS¹ dimension, the mass to charge range for scanning is set at 400-1400 - this will determine the parent ion spectrum. Secondly, the instrument has MS² capabilities whereby it will acquire fragmentation spectra of any parent ions whose intensities are detected to be greater than a predetermined threshold (Mann and Wilm, *Anal Chem* 66(24): 4390-4399 (1994)). A significant amount of information is collected for each protein sample as both a parent ion spectrum and many daughter ion spectra are generated with this instrumentation.

All resulting mass spectra are submitted to a database search algorithm for protein identification. A correlative mass algorithm is utilized along with a statistical verification of each match to identify a protein's identification (Ducret A, et al., *Protein Sci* 7(3): 706-719 (1998)). This method proves much more robust than MALD-TOF mass spectrometry for identifying the components of complex mixtures of proteins.

No interacting proteins were observed using at least one of the methods described above.

EXAMPLE 14 NMR Analysis

Purified protein sample is centrifuged at 13,000 rpm for 10 minutes with a bench-top microcentrifuge to eliminate any precipitated protein. The supernatant is then transferred into a clean tube and the sample volume is measured. If the sample volume is less than 450 µl, an appropriate amount of crystal buffer is added to the sample to reach that volume. Then 50 µl of D₂O (99.9%) is added to the sample to make an NMR sample of 500 µl. The usual concentration of the protein sample is usually approximately 1 mmol or greater.

NMR screening experiments are performed on a Bruker AV600 spectrometer equipped with a cryoprobe, or other equivalent instrumentation. All spectra are recorded at 25°C. Standard 1D proton pulse sequence with presaturation is used for 1D screening. Normally, a sweepwidth of 6400 Hz, and eight or sixteen scans are used, although different pulse sequences are known to those of skill in the art and may be readily determined. For ¹H, ¹⁵N HSQC experiments, a pulse sequence with "flip-back" water suppression may be used. Typically, sweepwidths of 8000 Hz and 2000 Hz are used for F2 and F1 dimension,

respectively. Four to sixteen scans are normally adequate. The data is then processed on a Sun Ultra 5 computer with NMRpipe software.

EXAMPLE 15 X-ray Crystallography

(a) Crystallization

Subsequent to purification, a subject polypeptide is centrifuged for 10 minutes at 4°C and at 14,000 rpm in order to sediment any aggregated protein. The protein sample is then diluted in order to provide multiple concentrations for screening.

Two 96 well plates (Nunc) are employed for the initial crystal screen, with 48 potential crystallization conditions. The screening library has crystallization conditions found in Hampton Research Crystal Screen I (Jankarik, J. and S.H. Kim, *J. Appl. Cryst.*, 1991. 24:409-11), Hampton Research Crystal Screen II, Hampton Crystal Screen I-Lite, and from Emerald Biostructures, Inc., Bainbridge Island, WA, Wizard I, Wizard II, Cryo I and Cryo II. Alternatively, other conditions known to those of skill in the art, including those provided in screening kits available from other companies, may also be tested.

Conditions are tested at multiple protein concentrations and at two temperatures (4 and 20°C). Crystal setups may be performed by a liquid handling robot appropriately programmed for sitting drop experiments. The robot loads 50 µl of buffer into each screening well on a 24 or 96 well sitting drop crystal screen tray, and then loads 1 - 5 µl of protein into each drop reservoir to be screened on the plate. Subsequently, the robot loads 1.5 µl of the corresponding screening solution into the drop reservoir atop the protein. The plate is then sealed using transparent tape, and stored at 4 or 20°C. Each plate is observed two days, two weeks, and 1 month after being set. Alternatively, screens may be performed using 0.1 - 10 µl drops suspended at the interface of two immiscible oils. The protein containing solution has a density intermediate between the two oils and thus floats between them (Chayen N.E.: 1996, *Protein Eng.* 9:927-29). This procedure may be performed in an automated fashion by an appropriately programmed liquid handling robot, with additional steps being required initially to introduce the oils. No tape is added to facilitate gradual drying out of the drop to promote crystallization.

Having identified conditions that are best suited for further crystal refinement, subsequent plates are set up to explore the affects of variables such as temperature, pH, salt

or PEG concentration on crystal size and form, with the intent of establishing conditions where the protein is able to form crystals of suitable size and morphology for diffraction analysis. Each refinement is performed in the sitting drop format in a 24 well Lindbro plate. Each well in the tray contains 500 μ l of screening solution, and a 1.5 μ l drop of protein diluted with 1.5 μ l of the screening solution is set to hang from the siliconized glass cover slip covering the well. Alternatively, refinement steps may be performed using either the machine 96 well plate hanging drop method or the oil suspension method described above.

Crystals of a selenomethionine-substituted polypeptide having the sequence of SEQ ID NO: 4, prepared and purified as described above and having a His tag, are obtained using the following conditions: 30% MPD, 0.1M sodium acetate pH 4.6, 20mM calcium chloride. In addition, crystals of the same polypeptide may be prepared under the following conditions: 0.4M potassium sodium tartrate. Further, crystals of the same polypeptide may be prepared under the following conditions: 30% PEG 4000, 0.1M TRIS pH 8.5, 0.2M magnesium chloride. Further, crystals of the same polypeptide may be prepared under the following conditions: 35% PEG 400, 0.2M Tris-HCl, pH 8.5, trisodium citrate dihydrate. The crystals were prepared using the following method: 20°C, sitting-drop, 10 mg/mL.

(b) Co-Crystallization

A variety of methods known in the art may be used for preparation of co-crystals comprising the subject polypeptides and one or more compounds that interact with the subject polypeptides, such as, for example, an inhibitor, co-factor, substrate, polynucleotide, polypeptide, and/or other molecule. In one exemplary method, crystals of the subject polypeptide may be soaked, for an appropriate period of time, in a solution containing a compound that interacts with a subject polypeptide. In another method, solutions of the subject polypeptide and/or compound that interacts with the subject polypeptide may be prepared for crystallization as described above and mixed into the above-described sitting drops. In certain embodiments, the molecule to be co-crystallized with the subject polypeptide may be present in the buffer in the sitting drop prior to addition of the solution comprising the subject polypeptide. In other embodiments, the subject polypeptide may be mixed with another molecule before adding the mixture to the sitting drop. Based on the teachings herein, one of skill in the art may determine the co-crystallization method yielding a co-crystal comprising the subject polypeptide.

(c) Heavy Atom Substitution

For preparation of crystals containing heavy atoms, crystals of the subject polypeptide may be soaked in a solution of a compound containing the appropriate heavy atom for such period as time as may be experimentally determined is necessary to obtain a useful heavy atom derivative for x-ray purposes. Likewise, for other compounds that may be of interest, including, for example, inhibitors or other molecules that interact with the subject polypeptide, crystals of the subject polypeptide may be soaked in a solution of such compound for an appropriate period of time.

(d) Data collection and processing

Crystals of TpiA (selenomethionyl) from *P. aeruginosa* were obtained at room temperature by the sitting drop vapour diffusion method in Tris-HCl (pH 8.5), 35% (vol/vol) PEG 400, and 0.2 M Trisodium Citrate Dihydrate. A suitable diamond-shaped crystal of TpiA (dimensions approx. 300 x 100 x 100 μ m) was selected for structure determination by single-crystal X-ray diffraction methods. The crystal in 20% ethylene glycol was frozen (100 K) in a nitrogen cold stream (Oxford Cryostream system) and diffraction data were collected at the Advanced Photon Source (Argonne National Laboratory) using the COM-CAT beamline equipped with a MAR CCD detector. The raw data were analyzed and reduced with the DENZO/SCALEPACK package. The protein was observed to crystallize in the hexagonal space group $P6_1$ with $a = 77.349$, $c = 175.528$ Å, and 2 molecules per asymmetric unit. Phase calculations, density modification and refinement were carried out using the CNX suite of programs. The coordinates from the structure of *E. coli* TIM (PDB ID #1TRE) were used as a starting model for phasing. Rigid body refinement was followed by refinement by maximum likelihood refinement against structure factors. The structure was partially built automatically by MAID, and completed manually by tracing the polypeptide chain in Turbo-Frodo. Refinement was performed in CNX, using simulated annealing torsion angle dynamics and individual B-factor refinement against maximum likelihood targets. Water molecules were automatically picked in CNX, and confirmed by visual inspection. A final inspection of $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ maps were used to locate all remaining ordered solvent molecules. Several iterations of model rebuilding in Turbo-Frodo and subsequent refinements were performed until convergence was achieved.

Structure solution and refined statistics are reported in Table 3, contained in FIGURE 8. FIGURE 9 contains a list of the atomic coordinates of the subject polypeptide and other molecules contained in the crystal. FIGURE 10 to FIGURE 16 depict various features of the crystal structure and other properties of a subject polypeptide.

(e) Analysis of the X-ray Structure of the Subject Polypeptide

General Description of Structure

P. aeruginosa triosephosphate isomerase (TIM) adopts a three-dimensional structure similar to other known TIM architectures. It is a homodimer consisting of two almost identical subunits. The monomer, consisting of 250 residues, is a single domain that can be described as an α/β barrel (or β barrel) due to the organization of its 8 $\alpha\beta$ units. The 8 β strands are aligned parallel to each other, forming a cylindrical inner structure (the "barrel") whose side is decorated by the adjoining α helices via the corresponding loops (FIGURE 11 and FIGURE 12). The common occurrence of this structure in numerous triosephosphate isomerases has resulted in the coinage of the more fitting description "TIM barrel". The geometric parameters of β barrels include the strand number ($n = 8$), the tilt of the β strands relative to the barrel axis ($\sim 36^\circ$), and the radius of the barrel ($\sim 6.5 - 7.5$ Å, depending on the eccentricity of the cross-section).

Active Site and Other Druggable Regions

Structurally, the active site appears to lie near the top of the TIM barrel towards the C-terminal ends of the β strands. The participating residues in the proton transfer reactions at the active site include the base Glu165 and acid His95 (amino acid numbering convention according to enzyme from chicken muscle). As shown in FIGURE 13 for *S. cerevisiae* TIM complexed with the transition state analogue 2-phosphoglycolytic acid, the other pertinent but non-catalytic residues at the active site appear to include Asn10 and Lys12, which may play a key role in anchoring the substrate at the catalytic site by binding its phosphate moiety via hydrogen bonds. Moreover, the positively charged side chain of Lys12 appears to help electrostatically stabilize the negatively charged transition state. The active site of *P. aeruginosa* TIM, comprising residues Glu165, His95, Asn10, and Lys12, and the surrounding residues, may present a druggable region.

During glycolysis, it is believed that the active site of TIM may accommodate either glyceraldehyde 3-phosphate or dihydroxyacetone phosphate, the interconversion of which

is thought to proceed via an unstable enediol intermediate. Thus, the enzyme appears to play an important role in both promoting the proton exchange to the form the enediol as well as stabilizing the intermediate. Analyses of TIM structures with and without inhibitory substrates bound have revealed an 11-residue (residues 166 – 176) flexible loop that, upon binding of a substrate, is believed to close over the active site. This conformational change involves main chain shifts of $> 7 \text{ \AA}$ as well as bringing the carboxylate group of the catalytic Glu165 approximately 2 \AA toward the substrate from its position in the unbound enzyme. During such proposed closure of the flexible loop, the hydroxyl group of the residue Tyr208 may form a hydrogen bond with the amide nitrogen of Ala176. Loop closure appears to be essential for shielding the active site from water and stabilizing the charged enediol transition state. The transition state, which normally decomposes in solution to form the toxic end product methylglyoxal, appears to be stabilized because the loop closure prevents the phosphate elimination reaction of the substrate. The open and closed conformations of the TIM structure are shown in FIGURE 14 and FIGURE 15.

The flexible loop appears to be a fairly conserved region of the enzyme, but variations have been observed between the yeast TIM sequence and the organisms of FIGURE 10. For example, position 175 is an alanine residue in yeast, but is a threonine residue in *P. aeruginosa*, and a serine residue in *E. coli*, *H. pylori*, *S. aureus*, *S. pneumoniae*, and *E. faecalis*. The loop in *P. aeruginosa* TIM enzyme comprising residues Pro166, Val167, Trp168, Ala169, Ile170, Gly171, Thr172, Gly173, Leu174, Thr175 and Ala176, may present a druggable region. The region closed off by and surrounding the loop may also may present a druggable region.

As shown in FIGURE 16, the active site involved in glycolysis in TIM structures is generally a conserved region of the enzyme. Thus, identifying inhibitors that bind to the active site of one TIM enzyme and not another may prove challenging. Other regions of the enzyme are under investigation for potential druggable regions that are more varied, including a region of significant differences between the trypanosomal (Ala100–Tyr101–Tyr102) and human (His100–Val101–Phe102) TIM enzymes, which is located approximately 15 \AA from the entrance to the true active site. In *P. aeruginosa* TIM enzyme, the catalytic center region is comprised of Leu100, Ile101, and Leu102, and may present a druggable region.

The structure of *T. cruzi* TIM (TcTIM) cocrystallized with hexane revealed three substrate molecules (H1, H2, H3) bound on the surface of TcTIM at regions that, in the

native structure, do not contain water molecules. One hexane molecule (H1) resides at < 4 Å from residues Arg135, Thr140, and Glu186 (part of loop 5 and helix 6) of subunit B. The other two, separated by 3.8 Å, are found in a hydrophobic patch composed of residues from both subunits: H2 is < 4 Å from Ile69, Tyr103, Gly104, Ile109, and Lys113 of subunit A and from Tyr102 and Tyr103 of subunit B. Similarly, H3 is found at < 4 Å from Phe75 (subunit A) and Tyr102 and Tyr103 (subunit B). The bound hexane molecules were observed to change the orientation of the side chain of the catalytic Glu residue compared to the native form, a long-range effect which could influence the catalytic activity of the complexed TIM. Of further interest is the proximity of H2 and H3 to the residues Ile69 and Phe75, which both belong to loop 3 (69 – 80) of subunit A. The regions in *P. aeruginosa* TIM that correspond to the first, second, and third hexane-binding regions are comprised of residues Arg134, Thr139, and Glu185 for the first hexane; Ala67, Leu102, Gly103, Val108, Lys112, Ile101, and Leu102 for the second hexane; and Leu74, Ile101, and Leu102 for the third hexane. These regions all present druggable regions.

It has been reported that the interactions of loop 3 from one subunit with residue 15 from the other subunit are important in dimer stability and enzyme catalysis. Selective inhibition of TcTIM Cys15 by sulfhydryl reagents at concentrations which do not affect human TIM (which has a Met at residue 15) or those lacking a cysteine in position 15 has been reported. Although in *P. aeruginosa* TIM this position is occupied by Met, the surrounding interdimer interface region of TIM nevertheless represents a druggable region, as there are 58 residues that participate in the intersubunit contacts of TIM. FIGURE 17 contains Table 4, which lists the residues in the intersubunit region of TcTIM (*T. cruzi* TIM), PaTIM (*P. aeruginosa* TIM), and HsTIM (*H. sapiens* TIM).

A comparison of the intersubunit residues in TcTIM, PaTIM, and humanTIM (listed in Table 4 of FIGURE 17) reveals that a unique Cys43 is found in PaTIM. This site may present a potential target residue for sulfhydryl agents or other drugs. There is also another unique cysteine (Cys88) which could be a potential target for such agents. FIGURE 18 depicts the location of these residues along the intersubunit interface. There are other residues that appear unique to PaTIM. For example, the polar, hydrophilic Gln48 is found in PaTIM, while in humanTIM the corresponding residue is the hydrophobic Phe50, and in TcTIM the nonpolar Met51. Likewise, a glutamine residue appears at position 82 in PaTIM, while the corresponding positions in TcTIM and human TIM are nonpolar, hydrophobic residues isoleucine and methionine, respectively. The differences in charge

and hydrophobicity in regions surrounding such distinct residues may be exploited in designing a drug that uniquely targets PaTIM. Thus, unique intersubunit residues and the area surrounding them may present druggable regions.

Known Inhibitors of Triosephosphate isomerase (TIM) Enzymes

Several known inhibitors of TIM enzymes are depicted in Figure 19:

Figure 19

| Inhibitor Name | Structure |
|------------------------------------|-----------|
| (R,S)-ornidazole | |
| (R,S)-alpha-chlorohydrin | |
| 1-chloro-3-hydroxypropanone | |
| D-glycerol 3-phosphate | |
| (S)-3-chlorolactaldehyde | |
| Bromoacetylanthanolamine phosphate | |

Other inhibitors include bromohydroxyacetone phosphate, glycidol phosphate, 2-phosphoglycolate, 2-phosphoglycolohydroxamate, 3-phosphopropionic acid, and 2-(N-formyl-N-hydroxy)-aminoethyl phosphonate.

The crystal structures of malarial TIM, *Trypanosoma cruzi*, *Plasmodium falciparum*, have been used in structure-based drug design efforts to discover antiparasitic agents. Since many parasites lack a functional tricarboxylic acid cycle, glycolysis is the sole energy source, making TIM an attractive target for antiprotozoal therapy.

Comparison to Other TIM enzymes

All TIM enzymes discovered to date have a TIM barrel structure. The TIM barrel consists of a cylinder composed of beta sheets surrounded by a rosette of alpha-helices. The catalytic mechanism of TIM action is still under active investigation, and the following discussion is not intended to be limiting in any way. These investigations have resulted in the following general findings: (1) TIM is active only as a dimer; (2) the structures of the two subunits are nearly identical, except at the two amino-terminal ends; (3) upon substrate binding, an 11-residue loop moves ~ 7 Å as a "lid" to close over the active site and the carboxylate group of the catalytic E165 moves ~ 2 Å toward the substrate from its position in the unbound enzyme. This motion has been deduced from many enzyme sources (chicken, trypanosome), as well as using various known enzyme inhibitors (phosphoglycolohydroxamate, 2-phosphoglycolate, 3-chloroacetol phosphate). The current hypothesis for the catalytic mechanism is an acid/base reaction in which E165 and H95 residues participate directly in proton transfer, in what is believed to be a concerted general acid – base catalysis involving low barrier hydrogen bonds. Other key residues are K12, Y208, and the 11-residue loop (aa 166-176). Wild-type triosephosphate isomerases appear to be active only as a dimer, although the enzyme is not cooperative.

Mutagenesis Studies

Site-directed mutagenesis of Glu165 to Asp165 in triosephosphate isomerase had the apparent effect of withdrawing the catalytic carboxylate group slightly away (~ 1 Å) from the substrate compared to its position in the native non-complexed conformation. The resulting deleterious effect on the activity of the mutant enzyme was reflected in the drastic decrease in the turnover (kcat) by 1000-fold. This experiment suggest that the glutamate residue in the active site is essential for full activity. Lys12 is believed to play a key role during substrate binding by fixing it in place in order for acid/base catalysis to occur.

When the Lys is replaced by Met in the K12M structure, the inhibitor PGH does not bind to the mutant isomerase, resulting in an open conformation of the enzyme

Based in part on the structural information described above, in one aspect, the present invention is directed towards druggable regions of a subject polypeptide or other TIM enzyme comprising the majority of the amino acid residues contained in a subject druggable region. In another aspect, the present invention is directed toward an inhibitor that interacts with the active site of such an enzyme. In another aspect, the present invention is directed towards an inhibitor that interacts with the flexible loop of such an enzyme so as to preclude it from closing, thereby inhibiting such enzyme. In still another aspect, the present invention is directed towards an inhibitor that interacts with the catalytic lid of such an enzyme so as to maintain the loop in the closed position. In yet another aspect, the present invention is directed towards an inhibitor that binds in the region that becomes closed off by the loop, and may optionally interact with the loop when bound. In another aspect, the present invention is directed towards the druggable region to which a hexane may bind in the subject structure. In another aspect, the present invention is directed toward the druggable regions along the interdimer interface.

EXAMPLE 16 Annotations

The functional annotation is arrived at by comparing the amino acid sequence of the ORF against all available ORFs in the NCBI database using BLAST. The closest match is selected to provide the probable function of the polypeptide having the sequence of SEQ ID NO: 2. Results of this comparison are described above and set forth in Table 2 of FIGURE 7.

The COGs database (Tatusov RL, Koonin EV, Lipman DJ. Science 1997; 278 (5338) 631-37) classifies proteins encoded in twenty-one completed genomes on the basis of sequence similarity. Members of the same Cluster of Orthologous Group, ("COG"), are expected to have the same or similar domain architecture and the same or substantially similar biological activity. The database may be used to predict the function of uncharacterised proteins through their homology to characterized proteins. The COGs database may be searched from NCBI's website (<http://www.ncbi.nlm.nih.gov/COG/>) to determine functional annotation descriptions, such as "information storage and processing" (translation, ribosomal structure and biogenesis, transcription, DNA replication,

recombination and repair); "cellular processes" (cell division and chromosome partitioning, post-translational modification, protein turnover, chaperones, cell envelope biogenesis, outer membrane, cell motility and secretion, inorganic ion transport and metabolism, signal transduction mechanisms); or "metabolism" (energy production and conversion, carbohydrate transport and metabolism, amino acid transport and metabolism, nucleotide transport and metabolism, coenzyme metabolism, lipid metabolism). For certain polypeptides, there is no entry available. Results of this analysis are described above and set forth in Table 2 of FIGURE 7.

EXAMPLE 17 Essential Gene Analysis

SEQ ID NO: 2 is compared to a number of publicly available "essential genes" lists to determine whether that protein is encoded by an essential gene. An example of such a list is descended from a free release at the www.shigen.nig.ac.jp PEC (profiling of *E. coli* chromosome) site, <http://www.shigen.nig.ac.jp/ecoli/pec/>. The list is prepared as follows: a wildcard search for all genes in class "essential" yields the list of essential *E. coli* proteins encoded by essential genes, which number 230. These 230 hits are pruned by comparing against an NCBI *E. coli* genome. Only 216 of the 230 genes on the list are found in the NCBI genome. These 216 are termed the essential-216-ecoli list. The essential-216-ecoli list is used to garner "essential" genes lists for other microbial genomes by blasting. For instance, formatting the 216-ecoli as a BLAST database, then BLASTing a genome (e.g. *S. aureus*) against it, elucidates all *S. aureus* genes with significant homology to a gene in the 216-essential list. SEQ ID NO: 2 is compared against the appropriate list and a match with a score of e^{-25} or better is considered an essential gene according to that list. In addition to the list described above, other lists of essential genes are publicly available or may be determined by methods disclosed publicly, and such lists and methods are considered in deciding whether a gene is essential. See, for example, Thanassi et al., *Nucleic Acids Res* 2002 Jul 15;30(14):3152-62; Forsyth et al., *Mol Microbiol* 2002 Mar;43(6):1387-400; Ji et al., *Science* 2001 Sep 21;293(5538):2266-9; Sassetti et al., *Proc Natl Acad Sci U S A* 2001 Oct 23;98(22):12712-7; Reich et al., *J Bacteriol* 1999 Aug;181(16):4961-8; Akerley et al., *Proc Natl Acad Sci U S A* 2002 Jan 22;99(2):966-71). Also, other methods are known in the art for determining whether a gene is essential, such as that disclosed in U.S. Patent Application No. 10/202,442 (filed July 24, 2002). The conclusion as to whether the gene

encoding the amino acid sequence set forth in SEQ ID NO: 2 is essential is set forth in Table 2 of FIGURE 7.

EXAMPLE 18 PDB Analysis

SEQ ID NO: 2 is compared against the amino acid sequences in a database of proteins whose structures have been solved and released to the PDB (protein data bank). The identity/information about the top PDB homolog (most similar "hit", if any; a PDB entry is only considered a hit if the score is e^{-4} or better) is annotated, and the percent similarity and identity between SEQ ID NO: 2 and the closest hit is calculated, with both being indicated in Table 2 of FIGURE 7.

EXAMPLE 19 Virtual Genome Analysis

VGDB or VG is a queryable collection of microbial genome databases annotated with biophysical and protein information. The organisms present in VG include:

| File | GRAM | Species | Source | Genome file date |
|-----------|------|-------------------------------|--------|-------------------|
| ecoli.faa | G- | <i>Escherichia coli</i> | NCBI | November 18 1998 |
| hpyl.faa | G- | <i>Helicobacter pylori</i> | NCBI | April 19 1999 |
| | | <i>Pseudomonas</i> | | |
| paer.faa | G- | <i>aeruginosa</i> | NCBI | September 22 2000 |
| ctra.faa | G- | <i>Chlamydia trachomatis</i> | NCBI | December 22 1999 |
| hinf.faa | G- | <i>Haemophilus influenzae</i> | NCBI | November 26 1999 |
| nmen.faa | G- | <i>Neisseria meningitidis</i> | NCBI | December 28 2000 |
| rpax.faa | G- | <i>Rickettsia prowazekii</i> | NCBI | December 22 1999 |
| bbur.faa | G- | <i>Borrelia burgdorferi</i> | NCBI | November 11 1998 |
| bsub.faa | G+ | <i>Bacillus subtilis</i> | NCBI | December 1 1999 |
| staph.faa | G+ | <i>Staphylococcus aureus</i> | TIGR | March 8 2001 |
| | | <i>Streptococcus</i> | | |
| spne.faa | G+ | <i>pneumoniae</i> | TIGR | February 22 2001 |
| mgen.faa | G+ | <i>Mycoplasma genitalium</i> | NCBI | November 23 1999 |
| efac.faa | G+ | <i>Enterococcus faecalis</i> | TIGR | March 8 2001 |

The VGDB comprises 13 microbial genomes, annotated with biophysical information (pI, MW, etc), and a wealth of other information. These 13 organism genomes are stored in a single flatfile (the VGDB) against which PSI-blast queries can be done.

SEQ ID NO: 2 is queried against the VGDB to determine whether this sequence is found, conserved, in many microbial genomes. There are certain criteria that must be met for a positive hit to be returned (beyond the criteria inherent in a basic PSI-blast).

When an ORF is queried it may have a maximum of 13 VG-organism hits. A hit is classified as such as long as it matches the following criteria: Minimum Length (as percentage of query length): 75 (*Ensure hit protein is at least 75% as long as query*); Maximum Length (as percentage of query length): 125 (*Ensure hit protein is no more than 125% as long as query*); eVal:-10 (*Ensure hit has an e-Value of e-10 or better*); Id%:>25 (*Ensure hit protein has at least 25% identity to query*). The e-Value is a standard parameter of BLAST sequence comparisons, and represents a measure of the similarity between two sequences based on the likelihood that any similarities between the two sequences could have occurred by random chance alone. The lower the e-Value, the less likely that the similarities could have occurred randomly and, generally, the more similar the two sequences are.

The organisms having an orthologue of the polypeptide having SEQ ID NO: 2 are listed in Table 2, shown in FIGURE 7.

EXAMPLE 20 Epitopic Regions

The three most likely epitopic regions of a polypeptide having SEQ ID NO: 2 are predicted using the semi-empirical method of Kolaskar and Tongaonkar (FEBS Letters 1990 v276 172-174), the software package called Protean (DNASTAR), or MacVectors's Protein analysis tools (Accelrys). The antigenic propensity of each amino acid is calculated by the ratio between frequency of occurrence of amino acids in 169 antigenic determinants experimentally determined and the calculated frequency of occurrence of amino acids at the surface of protein. The results of these bioinformatics analyses are presented in Table 2, shown in FIGURE 7.

EQUIVALENTS

The present invention provides among other things novel proteins, protein structures and protein-protein interactions. While specific embodiments of the subject invention have been discussed, the above specification is illustrative and not restrictive. Many variations

of the invention will become apparent to those skilled in the art upon review of this specification. The full scope of the invention should be determined by reference to the claims, along with their full scope of equivalents, and the specification, along with such variations.

All publications and patents mentioned herein, including those items listed below, are hereby incorporated by reference in their entirety as if each individual publication or patent was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control. To the extent that any U.S. Provisional Patent Applications to which this patent application claims priority incorporate by reference another U.S. Provisional Patent Application, such other U.S. Provisional Patent Application is not incorporated by reference herein unless this patent application expressly incorporates by reference, or claims priority to, such other U.S. Provisional Patent Application.

Also incorporated by reference in their entirety are any polynucleotide and polypeptide sequences which reference an accession number correlating to an entry in a public database, such as those maintained by The Institute for Genomic Research (TIGR) (www.tigr.org) and/or the National Center for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov).

Also incorporated by reference are the following: WO 00/45168, WO 00/79238, WO 00/77712, EP 1047108, EP 1047107, WO 00/72004, WO 00/73787, WO00/67017, WO 00/48004, WO 01/48209, WO 00/45168, WO 00/45164, U.S.S.N. 09/720272; PCT/CA99/00640; U.S. Patent Application Nos: 10/097125 (filed March 12, 2002); 10/097193 (filed March 12, 2002); 10/202442 (filed July 24, 2002); 10/097194 (filed March 12, 2002); 09/671817 (filed September 17, 2000); 09/965654 (filed September 27, 2001); 09/727812 (filed November 30, 2000); 60/370667 (filed April 8, 2002); a utility patent application entitled "Methods and Apparatuses for Purification" (filed September 18, 2002); U.S. Patent Numbers 6451591; 6254833; 6232114; 6229603; 6221612; 6214563; 6200762; 6171780; 6143492; 6124128; 6107477; D428157; 6063338; 6004808; 5985214; 5981200; 5928888; 5910287; 6248550; 6232114; 6229603; 6221612; 6214563; 6200762; 6197928; 6180411; 6171780; 6150176; 6140132; 6124128; 6107066; 6270988; 6077707; 6066476; 6063338; 6054321; 6054271; 6046925; 6031094; 6008378; 5998204; 5981200; 5955604; 5955453; 5948906; 5932474; 5925558; 5912137; 5910287; 5866548; 6214602; 5834436; 5777079; 5741657; 5693521; 5661035; 5625048; 5602258; 5552555; 5439797;

5374710; 5296703; 5283433; 5141627; 5134232; 5049673; 4806604; 4689432; 4603209; 6217873; 6174530; 6168784; 6271037; 6228654; 6184344; 6040133; 5910437; 5891993; 5854389; 5792664; 6248558; 6341256; 5854922; and 5866343; and Albin, R. and P. M. Silverman (1984). *Mol Gen Genet* 197(2): 261-71; Aqvist, J. and M. Fothergill (1996). *J Biol Chem* 271(17): 10010-6; Burton, P. M. and S. G. Waley (1968). *Biochim Biophys Acta* 151(3): 714-5; Campbell, I. D., R. B. Jones, et al. (1979). *Biochem J* 179(3): 607-21; Delboni, L. F., S. C. Mande, et al. (1995). *Protein Sci* 4(12): 2594-604; Fenn, R. H. and G. E. Marshall (1972). *Biochem J* 130(1): 1-10; Garza-Ramos, G., N. Cabrera, et al. (1998). *Eur J Biochem* 253(3): 684-91; Gibson, D. R., R. W. Gracy, et al. (1980). *J Biol Chem* 255(19): 9369-74; Gomez-Puyou, A., E. Saavedra-Lira, et al. (1995). *Chem Biol* 2(12): 847-55; Hartman, F. C. (1968). *Biochem Biophys Res Commun* 33(6): 888-94; Hartman, F. C. (1970). *Biochemistry* 9(8): 1783-91; Hartman, F. C. (1970). *Biochemistry* 9(8): 1776-82; Hartman, F. C. (1971). *Biochemistry* 10(1): 146-54; Hartman, F. C., G. M. LaMuraglia, et al. (1975). *Biochemistry* 14(24): 5274-9; Hartman, F. C. and I. C. Norton (1977). *Methods Enzymol* 47: 479-98; Heinz, D. W., M. Ryan, et al. (1995). *Embo J* 14(16): 3855-63; Johnson, L. N. and R. Wolfenden (1970). *J Mol Biol* 47(1): 93-100; Jones, R. B. and S. G. Waley (1979). *Biochem J* 179(3): 623-30; Jones, A. R. and S. J. Cooney (1987). *Biochem Biophys Res Commun* 145(3): 1054-8; Jones, A. R. and L. M. Porter (1995). *Reprod Fertil Dev* 7(5): 1089-94; Joubert, F., A. W. Neitz, et al. (2001). *Proteins* 45(2): 136-43; Krietsch, W. K., P. G. Pentchev, et al. (1970). *Eur J Biochem* 14(2): 289-300; Kursula, I., S. Partanen, et al. (2001). *Eur J Biochem* 268(19): 5189-96; Lolis, E. and G. A. Petsko (1990). *Biochemistry* 29(28): 6619-25; Marks, G. T., T. K. Harris, et al. (2001). *Biochemistry* 40(23): 6805-18; Mendz, G. L., S. L. Hazell, et al. (1994). *Arch Biochem Biophys* 312(2): 349-56; Nader, W., A. Betz, et al. (1979). *Biochim Biophys Acta* 571(2): 177-85; Niitsu, Y., O. Hori, et al. (1999). *Brain Res Mol Brain Res* 74(1-2): 26-34; Noble, M. E., C. L. Verlinde, et al. (1991). *J Med Chem* 34(9): 2709-18; Noble, M. E., R. K. Wierenga, et al. (1991). *Proteins* 10(1): 50-69; Norton, I. L. and F. C. Hartman (1972). *Biochemistry* 11(24): 4435-41; O'Connell, E. L. and I. A. Rose (1977). *Methods Enzymol* 46: 381-8; Ostoa-Saloma, P., G. Garza-Ramos, et al. (1997). *Eur J Biochem* 244(3): 700-5; Rose, I. A. and E. L. O'Connell (1969). *J Biol Chem* 244(23): 6548-50; Saadat, D. and D. H. Harrison (2000). *Biochemistry* 39(11): 2950-60; Thomas, M. K. and T. G. Spring (1976). *Biochem J* 153(3): 741-4; Verlinde, C. L. M. J.; Rudenko, G.; Hol, W. G. J. *J. Comput. Aided Mol. Design* 6, 131 (1992); Gao, X.-G., et al. (1999) *Proc. Natl. Acad. Sci. USA* 96, 10062;

Maldonado, E., et al. (1998) *J. Mol. Biol.* 283, 193 ; Raines, R. T., et al. (1986) *Biochemistry* 25, 7142-7154; Joseph-McCarthy, D. D., et al. (1994) *Biochem.* 33, 2815; Otwinowski, Z. and W. Minor (1997). *Methods in Enzymology*, vol. 276: *Macromolecular Crystallography (Part A)*, p. 307-326, C. W. Carter, Jr. and R. M. Sweet, Eds., Academic Press; Brunger, A. T. et al. (1998) *Acta Crystallogr. D. Biol. Crystallogr.* 54, 905-921; Noble, M. E. M., et al, J. A. (1993) *Acta Crystallogr D Biol Crystallogr* 49 403; Velanker, S. S., S. S. Ray, et al. (1997). *Structure* 5(6): 751-61; Verlinde, C. L., C. J. Witmans, et al. (1992). *Protein Sci* 1(12): 1578-84; Verlinde, C. L., G. Rudenko, et al. (1992). *J Comput Aided Mol Des* 6(2): 131-47; Verlinde, C. L., E. A. Merritt, et al. (1994). *Protein Sci* 3(10): 1670-86; E.Lolis, G.A.Petsko. *Biochem.* 29, 6619 (1990); Waley, S. G., J. C. Miller, et al. (1970) *Nature* 227(254): 1811 and Zubillaga, R. A., R. Perez-Montfort, et al. (1994). *Arch Biochem Biophys* 313(2): 328-36.

We claim:

CLAIMS

1. A composition comprising an isolated, recombinant polypeptide, wherein the polypeptide comprises: (a) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (b) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (c) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; and wherein the polypeptide of (a), (b) or (c) is at least about 90% pure in a sample of the composition.

2. The composition of claim 1, wherein the polypeptide is at least about 95% pure as determined by gel electrophoresis.

3. The composition of claim 1, wherein the polypeptide is purified to essential homogeneity.

4. The composition of claim 1, wherein at least about two-thirds of the polypeptide in the sample is soluble.

5. The composition of claim 1, wherein the polypeptide is fused to at least one heterologous polypeptide that increases the solubility or stability of the polypeptide.

6. The composition of claim 1, which further comprises a matrix suitable for mass spectrometry.

7. The composition of claim 6, wherein the matrix is a nicotinic acid derivative or a cinnamic acid derivative.

8. A sample comprising an isolated, recombinant polypeptide, wherein the polypeptide comprises: (a) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (b) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (c) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; and wherein the polypeptide of (a), (b) or (c) is labeled with a heavy atom.

9. The sample of claim 8, wherein the heavy atom is one of the following: cobalt, selenium, krypton, bromine, strontium, molybdenum, ruthenium, rhodium, palladium, silver, cadmium, tin, iodine, xenon, barium, lanthanum, cerium, praseodymium, neodymium, samarium, europium, gadolinium, terbium, dysprosium, holmium, erbium, thulium, ytterbium, lutetium, tantalum, tungsten, rhenium, osmium, iridium, platinum, gold, mercury, thallium, lead, thorium and uranium.

10. The sample of claim 8, wherein the polypeptide is labeled with seleno-methionine.

11. The sample of claim 8, further comprising a cryo-protectant.

12. The sample of claim 11, wherein the cryo-protectant is one of the following: methyl pentanediol, isopropanol, ethylene glycol, glycerol, formate, citrate, mineral oil and a low-molecular-weight polyethylene glycol.

13. A crystallized, recombinant polypeptide comprising: (a) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (b) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (c) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; wherein the polypeptide of (a), (b) or (c) is in crystal form.

14. A crystallized complex comprising the crystallized, recombinant polypeptide of claim 13 and a co-factor, wherein the complex is in crystal form.

15. A crystallized complex comprising the crystallized, recombinant polypeptide of claim 13 and a small organic molecule, wherein the complex is in crystal form.

16. The crystallized, recombinant polypeptide of claim 13, which diffracts x-rays to a resolution of about 3.5 Å or better.

17. The crystallized, recombinant polypeptide of claim 13, wherein the polypeptide comprises at least one heavy atom label.

18. The crystallized, recombinant polypeptide of claim 17, wherein the polypeptide is labeled with seleno-methionine.

19. A method for designing a modulator for the prevention or treatment of *P. aeruginosa* related disease or disorder, comprising:

(a) providing a three-dimensional structure for a crystallized, recombinant polypeptide of claim 13;

(b) identifying a potential modulator for the prevention or treatment of *P. aeruginosa* related disease or disorder by reference to the three-dimensional structure;

(c) contacting a polypeptide of the composition of claim 1 or *P. aeruginosa* with the potential modulator; and

(d) assaying the activity of the polypeptide or determining the viability of *P. aeruginosa* after contact with the modulator, wherein a change in the activity of the polypeptide or the viability of *P. aeruginosa* indicates that the modulator may be useful for prevention or treatment of a *P. aeruginosa* related disease or disorder.

20. A sample comprising an isolated, recombinant polypeptide, wherein the polypeptide comprises: (a) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (b) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (c) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; and wherein the polypeptide of (a), (b) or (c) is enriched in at least one NMR isotope.

21. The sample of claim 20, wherein the NMR isotope is one of the following: hydrogen-1 (^1H), hydrogen-2 (^2H), hydrogen-3 (^3H), phosphorous-31 (^{31}P), sodium-23 (^{23}Na), nitrogen-14 (^{14}N), nitrogen-15 (^{15}N), carbon-13 (^{13}C) and fluorine-19 (^{19}F).

22. The sample of claim 20, further comprising a deuterium lock solvent.

23. The sample of claim 22, wherein the deuterium lock solvent is one of the following: acetone (CD_3COCD_3), chloroform (CDCl_3), dichloro methane (CD_2Cl_2), methylnitrile (CD_3CN), benzene (C_6D_6), water (D_2O), diethylether ($(\text{CD}_3\text{CD}_2)_2\text{O}$), dimethylether ($(\text{CD}_3)_2\text{O}$), N,N-dimethylformamide ($(\text{CD}_3)_2\text{NCDO}$), dimethyl sulfoxide (CD_3SOCD_3), ethanol ($\text{CD}_3\text{CD}_2\text{OD}$), methanol (CD_3OD), tetrahydrofuran ($\text{C}_4\text{D}_8\text{O}$), toluene ($\text{C}_6\text{D}_5\text{CD}_3$), pyridine ($\text{C}_5\text{D}_5\text{N}$) and cyclohexane (C_6H_{12}).

24. The sample of claim 20, which is contained within an NMR tube.

25. A method for identifying small molecules that bind to a polypeptide of the composition of claim 1, comprising:

(a) generating a first NMR spectrum of an isotopically labeled polypeptide of the composition of claim 1;

(b) exposing the polypeptide to one or more small molecules;

(c) generating a second NMR spectrum of the polypeptide which has been exposed to one or more small molecules; and

(d) comparing the first and second spectra to determine differences between the first and the second spectra, wherein the differences are indicative of one or more small molecules that have bound to the polypeptide.

26. A host cell comprising a nucleic acid encoding a polypeptide comprising: (a) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (b) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (c) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; wherein a culture of the host cell produces at least about 1 mg of the polypeptide per liter of culture and the polypeptide is at least about one-third soluble as measured by gel electrophoresis.

27. An isolated, recombinant polypeptide, comprising: (a) an amino acid sequence having at least about 90% identity with the amino acid sequence set forth in SEQ ID NO: 4; or (b) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; and wherein the polypeptide comprises one or more of the following amino acid residues at the specified position of the polypeptide: N at position 152, and S at position 233.

28. A method for obtaining structural information of a crystallized polypeptide, the method comprising:

(a) crystallizing a recombinant polypeptide, wherein the polypeptide comprises: (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide

that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; and wherein the crystallized polypeptide is capable of diffracting X-rays to a resolution of 3.5 Å or better; and

(b) analyzing the crystallized polypeptide by X-ray diffraction to determine the three-dimensional structure of at least a portion of the crystallized polypeptide.

29. The method of claim 28, wherein the three-dimensional structure of the portion of the crystallized polypeptide is determined to a resolution of 3.5 Å or better.

30. A method for identifying a druggable region of a polypeptide, the method comprising:

(a) obtaining crystals of a polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*, such that the three dimensional structure of the crystallized polypeptide may be determined to a resolution of 3.5 Å or better;

(b) determining the three dimensional structure of the crystallized polypeptide using X-ray diffraction; and

(c) identifying a druggable region of the crystallized polypeptide based on the three-dimensional structure of the crystallized polypeptide.

31. The method of claim 30, wherein the druggable region is an active site.

32. The method of claim 31, wherein the druggable region is on the surface of the polypeptide.

33. Crystalline triosephosphate isomerase from *P. aeruginosa* comprising a hexagonal crystal having unit cell dimensions of $a = 77.349$, $c = 175.528$ Å, and space group $P6_1$, the unit cell containing two molecules per asymmetric unit.

34. A crystallized polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3)

an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; wherein the crystal has a P6₁ space group.

35. A crystallized polypeptide comprising a structure of a polypeptide that is defined by a substantial portion of the atomic coordinates set forth in FIGURE 9.

36. A method for determining the crystal structure of a homolog of a polypeptide, the method comprising:

(a) providing the three dimensional structure of a first crystallized polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*;

(b) obtaining crystals of a second polypeptide comprising an amino acid sequence that is at least 70% identical to the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4, such that the three dimensional structure of the second crystallized polypeptide may be determined to a resolution of 3.5 Å or better; and

(c) determining the three dimensional structure of the second crystallized polypeptide by x-ray crystallography based on the atomic coordinates of the three dimensional structure provided in step (a).

37. The method of claim 36, wherein the atomic coordinates for the second crystallized polypeptide have a root mean square deviation from the backbone atoms of the first polypeptide of not more than 1.5 Å for all backbone atoms shared in common with the first polypeptide and the second polypeptide.

38. A method for homology modeling a homolog of triosephosphate isomerase from *P. aeruginosa*, comprising:

(a) aligning the amino acid sequence of a homolog of triosephosphate isomerase from *P. aeruginosa* with an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 and incorporating the sequence of the homolog of triosephosphate isomerase from

P. aeruginosa into a model of triosephosphate isomerase from *P. aeruginosa* derived from structure coordinates as listed in FIGURE 9 to yield a preliminary model of the homolog of triosephosphate isomerase from *P. aeruginosa*;

(b) subjecting the preliminary model to energy minimization to yield an energy minimized model;

(c) remodeling regions of the energy minimized model where stereochemistry restraints are violated to yield a final model of the homolog of triosephosphate isomerase from *P. aeruginosa*.

39. A method for obtaining structural information about a molecule or a molecular complex of unknown structure comprising:

(a) crystallizing the molecule or molecular complex;

(b) generating an x-ray diffraction pattern from the crystallized molecule or molecular complex;

(c) applying at least a portion of the structure coordinates set forth in FIGURE 9 to the x-ray diffraction pattern to generate a three-dimensional electron density map of at least a portion of the molecule or molecular complex whose structure is unknown.

40. A method for attempting to make a crystallized complex comprising a polypeptide and a modulator having a molecular weight of less than 5 kDa, the method comprising:

(a) crystallizing a polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; such that crystals of the crystallized polypeptide will diffract x-rays to a resolution of 5 Å or better; and

(b) soaking the crystals in a solution comprising a potential modulator having a molecular weight of less than 5 kDa.

41. A method for incorporating a potential modulator in a crystal of a polypeptide, comprising placing a hexagonal crystal of triosephosphate isomerase from *P. aeruginosa*

having unit cell dimensions of $a = 77.349$, $c = 175.528$ Å and space group $P6_1$ in a solution comprising the potential modulator.

42. A computer readable storage medium comprising digitally encoded structural data, wherein the data comprises structural coordinates as listed in FIGURE 9 for the backbone atoms of at least about six amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa*.

43. A scalable three-dimensional configuration of points, at least a portion of the points derived from some or all of the structure coordinates as listed in FIGURE 9 for a plurality of amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa*.

44. The scalable three-dimensional configuration of points of claim 43, wherein the structure coordinates as listed in FIGURE 9 for the backbone atoms of at least about five amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa* are used to derive part or all of the portion of points.

45. The scalable three-dimensional configuration of points of claim 43, wherein the structure coordinates as listed in FIGURE 9 for the backbone and optionally the side chain atoms of at least about ten amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa* are used to derive part or all of the portion of points.

46. The scalable three-dimensional configuration of points of claim 43, wherein the structure coordinates as listed in FIGURE 9 for the backbone atoms of at least about fifteen amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa* are used to derive part or all of the portion of points.

47. The scalable three-dimensional configuration of points of claim 43, wherein substantially all of the points are derived from structure coordinates as listed in FIGURE 9.

48. The scalable three-dimensional configuration of points of claim 43, wherein the structure coordinates as listed in FIGURE 9 for one or more of the following groups of atoms from a druggable region of triosephosphate isomerase from *P. aeruginosa* are used to derive part or all of the portion of points:

(i) the backbone atoms of at least a majority of Pro166, Val167, Trp168, Ala169, Ile170, Gly171, Thr172, Gly173, Leu174, Thr175 and Ala176;

(ii) the backbone atoms of Glu165 and the backbone atoms and three or more of the side chain atoms of His95;

(iii) the backbone atoms of Glu165, His95, Asn10 and Lys12;

(iv) the backbone atoms for all of Arg134, Thr139, and Glu185;

(v) the backbone atoms for at least four of Ala67, Leu102, Gly103, Val108, Lys112, Ile101, and Leu102;

(vi) the backbone atoms for all of Leu74, Ile101 and Leu102;

(vii) the backbone atoms of at least about ten amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa*;

(viii) the backbone atoms of at least ten of the amino acid residues identified in Table 4 of FIGURE 17; or

(ix) the backbone and side chain atoms of Cys43 and optionally Cys88 and the backbone atoms of at least three other amino acid residues that may be involved in the intersubunit interaction of a dimer of triosephosphate isomerase from *P. aeruginosa*.

49. A scalable three-dimensional configuration of points, comprising points having a root mean square deviation of less than about 1.5 Å from the three dimensional coordinates as listed in FIGURE 9 for the backbone atoms of at least five amino acid residues, wherein the five amino acid residues are from a druggable region of triosephosphate isomerase from *P. aeruginosa*.

50. The scalable three-dimensional configuration of points of claim 49, wherein any point-to-point distance, calculated from the three dimensional coordinates as listed in FIGURE 9, between one of the backbone atoms for one of the five amino acid residues and another backbone atom of a different one of the five amino acid residues is not more than about 10 Å.

51. A scalable three-dimensional configuration of points comprising points having a root mean square deviation of less than about 1.5 Å from the three dimensional coordinates as listed in FIGURE 9 for one or more of the following groups of atoms from a druggable region of triosephosphate isomerase from *P. aeruginosa*:

(i) the backbone atoms of at least a majority of Pro166, Val167, Trp168, Ala169, Ile170, Gly171, Thr172, Gly173, Leu174, Thr175 and Ala176;

(ii) the backbone atoms of Glu165 and the backbone atoms and three or more of the side chain atoms of His95;

(iii) the backbone atoms of Glu165, His95, Asn10 and Lys12;

(iv) the backbone atoms for all of Arg134, Thr139, and Glu185;

(v) the backbone atoms for at least four of Ala67, Leu102, Gly103, Val108, Lys112, Ile101, and Leu102;

(vi) the backbone atoms for all of Leu74, Ile101 and Leu102;

(vii) the backbone atoms of at least about ten amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa*;

(viii) the backbone atoms of at least ten of the amino acid residues identified in Table 4 of FIGURE 17; or

(ix) the backbone and side chain atoms of Cys43 and optionally Cys88 and the backbone atoms of at least three other amino acid residues that may be involved in the intersubunit interaction of a dimer of triosephosphate isomerase from *P. aeruginosa*.

52. A computer readable storage medium comprising digitally encoded structural data, wherein the data comprise the identity and three-dimensional coordinates as listed in FIGURE 9 for one of the following groups of atoms from a druggable region of triosephosphate isomerase from *P. aeruginosa*:

(i) the backbone atoms of at least a majority of Pro166, Val167, Trp168, Ala169, Ile170, Gly171, Thr172, Gly173, Leu174, Thr175 and Ala176;

(ii) the backbone atoms of Glu165 and the backbone atoms and three or more of the side chain atoms of His95;

(iii) the backbone atoms of Glu165, His95, Asn10 and Lys12;

(iv) the backbone atoms for all of Arg134, Thr139, and Glu185;

(v) the backbone atoms for at least four of Ala67, Leu102, Gly103, Val108, Lys112, Ile101, and Leu102;

(vi) the backbone atoms for all of Leu74, Ile101 and Leu102;

(vii) the backbone atoms of at least about ten amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa*;

(viii) the backbone atoms of at least ten of the amino acid residues identified in Table 4 of FIGURE 17; or

(ix) the backbone and side chain atoms of Cys43 and optionally Cys88 and the backbone atoms of at least three other amino acid residues that may be involved in the intersubunit interaction of a dimer of triosephosphate isomerase from *P. aeruginosa*.

53. A scalable three-dimensional configuration of points, wherein the points have a root mean square deviation of less than about 1.5 Å from the three dimensional coordinates as listed in FIGURE 9 for the atoms of one or more of the following groups from a druggable region of triosephosphate isomerase from *P. aeruginosa*:

(i) the atoms of at least a majority of Pro166, Val167, Trp168, Ala169, Ile170, Gly171, Thr172, Gly173, Leu174, Thr175 and Ala176;

(ii) the atoms of Glu165 and His95;

(iii) the atoms of Glu165, His95, Asn10 and Lys12;

(iv) the atoms for all of Arg134, Thr139, and Glu185;

(v) the atoms for at least four of Ala67, Leu102, Gly103, Val108, Lys112, Ile101, and Leu102;

(vi) the atoms for all of Leu74, Ile101 and Leu102;

(vii) the atoms of at least about ten amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa*;

(viii) the atoms of at least ten of the amino acid residues identified in Table 4 of FIGURE 17; or

(ix) the atoms of Cys43 and optionally Cys88 and the atoms of at least three other amino acid residues that may be involved in the intersubunit interaction of a dimer of triosephosphate isomerase from *P. aeruginosa*;

wherein up to one amino acid residue in each of the groups (i) to (ix) may have a conservative substitution thereof.

54. A scalable three-dimensional configuration of points derived from a druggable region of a polypeptide, wherein the points have a root mean square deviation of less than about 1.5 Å from the three dimensional coordinates as listed in FIGURE 9 for the backbone

atoms of at least ten amino acid residues that participate in the intersubunit contacts of triosephosphate isomerase from *P. aeruginosa*.

55. The scalable three-dimensional configuration of points of claim 54, wherein one of the ten amino acid residues is Cys43.

56. A computer-assisted method for identifying an inhibitor of the activity of triosephosphate isomerase from *P. aeruginosa*, comprising:

(a) supplying a computer modeling application with a set of structure coordinates as listed in FIGURE 9 for atoms from one or more of the following groups so as to define part or all of a molecule or complex:

(i) the backbone atoms of at least a majority of Pro166, Val167, Trp168, Ala169, Ile170, Gly171, Thr172, Gly173, Leu174, Thr175 and Ala176;

(ii) the backbone atoms of Glu165 and the backbone atoms and three or more of the side chain atoms of His95;

(iii) the backbone atoms of Glu165, His95, Asn10 and Lys12;

(iv) the backbone atoms for all of Arg134, Thr139, and Glu185;

(v) the backbone atoms for at least four of Ala67, Leu102, Gly103, Val108, Lys112, Ile101, and Leu102;

(vi) the backbone atoms for all of Leu74, Ile101 and Leu102;

(vii) the backbone atoms of at least about ten amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa*;

(viii) the backbone atoms of at least ten of the amino acid residues identified in Table 4 of FIGURE 17; or

(ix) the backbone and side chain atoms of Cys43 and optionally Cys88 and the backbone atoms of at least three other amino acid residues that may be involved in the intersubunit interaction of a dimer of triosephosphate isomerase from *P. aeruginosa*;

(b) supplying the computer modeling application with a set of structure coordinates of a chemical entity; and

(c) determining whether the chemical entity is expected to bind to or interfere with the molecule or complex.

57. The method of claim 56, wherein determining whether the chemical entity is expected to bind to or interfere with the molecule or complex comprises performing a fitting operation between the chemical entity and a druggable region of the molecule or complex, followed by computationally analyzing the results of the fitting operation to quantify the association between the chemical entity and the druggable region.

58. The method of claim 56, further comprising screening a library of chemical entities.

59. A computer-assisted method for designing an inhibitor of triosephosphate isomerase activity comprising:

(a) supplying a computer modeling application with a set of structure coordinates having a root mean square deviation of less than about 1.5 Å from the structure coordinates as listed in FIGURE 9 for atoms from one or more of the following groups so as to define part or all of a molecule or complex:

(i) the backbone atoms of at least a majority of Pro166, Val167, Trp168, Ala169, Ile170, Gly171, Thr172, Gly173, Leu174, Thr175 and Ala176;

(ii) the backbone atoms of Glu165 and the backbone atoms and three or more of the side chain atoms of His95;

(iii) the backbone atoms of Glu165, His95, Asn10 and Lys12;

(iv) the backbone atoms for all of Arg134, Thr139, and Glu185;

(v) the backbone atoms for at least four of Ala67, Leu102, Gly103, Val108, Lys112, Ile101, and Leu102;

(vi) the backbone atoms for all of Leu74, Ile101 and Leu102;

(vii) the backbone atoms of at least about ten amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa*;

(viii) the backbone atoms of at least ten of the amino acid residues identified in Table 4 of FIGURE 17; or

(ix) the backbone and side chain atoms of Cys43 and optionally Cys88 and the backbone atoms of at least three other amino acid residues that may be involved in the intersubunit interaction of a dimer of triosephosphate isomerase from *P. aeruginosa*;

(b) supplying the computer modeling application with a set of structure coordinates for a chemical entity;

(c) evaluating the potential binding interactions between the chemical entity and the molecule or complex;

(d) structurally modifying the chemical entity to yield a set of structure coordinates for a modified chemical entity; and

(e) determining whether the modified chemical entity is an inhibitor expected to bind to or interfere with the molecule or complex, wherein binding to or interfering with the molecule or molecular complex is indicative of potential inhibition of triosephosphate isomerase activity.

60. The method of claim 59, wherein determining whether the modified chemical entity is an inhibitor expected to bind to or interfere with the molecule or complex comprises performing a fitting operation between the chemical entity and the molecule or complex, followed by computationally analyzing the results of the fitting operation to evaluate the association between the chemical entity and the molecule or complex.

61. The method of claim 59, wherein the set of structure coordinates for the chemical entity is obtained from a chemical library.

62. A computer-assisted method for designing an inhibitor of triosephosphate isomerase activity *de novo* comprising:

(a) supplying a computer modeling application with a set of three-dimensional coordinates derived from the structure coordinates as listed in FIGURE 9 for atoms from one or more of the following groups so as to define part or all of a molecule or complex:

(i) the backbone atoms of at least a majority of Pro166, Val167, Trp168, Ala169, Ile170, Gly171, Thr172, Gly173, Leu174, Thr175 and Ala176;

(ii) the backbone atoms of Glu165 and the backbone atoms and three or more of the side chain atoms of His95;

(iii) the backbone atoms of Glu165, His95, Asn10 and Lys12;

(iv) the backbone atoms for all of Arg134, Thr139, and Glu185;

(v) the backbone atoms for at least four of Ala67, Leu102, Gly103, Val108, Lys112, Ile101, and Leu102;

- (vi) the backbone atoms for all of Leu74, Ile101 and Leu102;
 - (vii) the backbone atoms of at least about ten amino acid residues from a druggable region of triosephosphate isomerase from *P. aeruginosa*;
 - (viii) the backbone atoms of at least ten of the amino acid residues identified in Table 4 of FIGURE 17; or
 - (ix) the backbone and side chain atoms of Cys43 and optionally Cys88 and the backbone atoms of at least three other amino acid residues that may be involved in the intersubunit interaction of a dimer of triosephosphate isomerase from *P. aeruginosa*;
- (b) computationally building a chemical entity represented by a set of structure coordinates; and
- (c) determining whether the chemical entity is an inhibitor expected to bind to or interfere with the molecule or complex, wherein binding to or interfering with the molecule or complex is indicative of potential inhibition of triosephosphate isomerase activity.

63. The method of claim 62, wherein determining whether the chemical entity is an inhibitor expected to bind to or interfere with the molecule or complex comprises performing a fitting operation between the chemical entity and a druggable region of the molecule or complex, followed by computationally analyzing the results of the fitting operation to quantify the association between the chemical entity and the druggable region.

64. The method of any of claims 56, 59 or 62, further comprising supplying or synthesizing the potential inhibitor, then assaying the potential inhibitor to determine whether it inhibits triosephosphate isomerase activity.

65. A method for identifying a potential modulator for the prevention or treatment of a *P. aeruginosa* related disease or disorder, the method comprising:

- (a) providing the three dimensional structure of a crystallized polypeptide comprising: (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*;

(b) obtaining a potential modulator for the prevention or treatment of *P. aeruginosa* related disease or disorder based on the three dimensional structure of the crystallized polypeptide;

(c) contacting the potential modulator with a second polypeptide comprising: (i) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (ii) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (iii) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; which second polypeptide may optionally be the same as the crystallized polypeptide; and

(d) assaying the activity of the second polypeptide, wherein a change in the activity of the second polypeptide indicates that the compound may be useful for prevention or treatment of a *P. aeruginosa* related disease or disorder.

66. A method for designing a candidate modulator for screening for inhibitors of a polypeptide, the method comprising:

(a) providing the three dimensional structure of a druggable region of a polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*; and

(b) designing a candidate modulator based on the three dimensional structure of the druggable region of the polypeptide.

67. A method for identifying a potential modulator of a polypeptide from a database, the method comprising:

(a) providing the three-dimensional coordinates for a plurality of the amino acids of a polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence

encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*;

(b) identifying a druggable region of the polypeptide; and

(c) selecting from a database at least one potential modulator comprising three dimensional coordinates which indicate that the modulator may bind or interfere with the druggable region.

68. The method of claim 67, wherein the modulator is a small molecule.

69. A method for preparing a potential modulator of a druggable region contained in a polypeptide, the method comprising:

(a) using the atomic coordinates for the backbone atoms of at least about six amino acid residues from a polypeptide of SEQ ID NO: 4, with a \pm a root mean square deviation from the backbone atoms of the amino acid residues of not more than 1.5 Å, to generate one or more three-dimensional structures of a molecule comprising a druggable region from the polypeptide;

(b) employing one or more of the three dimensional structures of the molecule to design or select a potential modulator of the druggable region; and

(c) synthesizing or obtaining the modulator.

70. An apparatus for determining whether a compound is a potential modulator of a polypeptide, the apparatus comprising:

(a) a memory that comprises:

(i) the three dimensional coordinates and identities of at least about fifteen atoms from a druggable region of a polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of triosephosphate isomerase from *P. aeruginosa*;

(ii) executable instructions; and

- (b) a processor that is capable of executing instructions to:
- (i) receive three-dimensional structural information for a candidate modulator;
 - (ii) determine if the three-dimensional structure of the candidate modulator is complementary to the three dimensional coordinates of the atoms from the druggable region; and
 - (iii) output the results of the determination.

71. A method for making an inhibitor of triosephosphate isomerase activity, the method comprising chemically or enzymatically synthesizing a chemical entity to yield an inhibitor of triosephosphate isomerase activity, the chemical entity having been identified during a computer-assisted process comprising supplying a computer modeling application with a set of structure coordinates of a molecule or complex, the molecule or complex comprising at least a portion of at least one druggable region from triosephosphate isomerase from *P. aeruginosa*; supplying the computer modeling application with a set of structure coordinates of a chemical entity; and determining whether the chemical entity is expected to bind or to interfere with the molecule or complex at a druggable region, wherein binding to or interfering with the molecule or complex is indicative of potential inhibition of triosephosphate isomerase activity.

72. A computer readable storage medium comprising digitally encoded data, wherein the data comprises structural coordinates for a druggable region that is structurally homologous to the structure coordinates as listed in FIGURE 9 for a druggable region of triosephosphate isomerase from *P. aeruginosa*.

73. A computer readable storage medium comprising digitally encoded structural data, wherein the data comprise a majority of the three-dimensional structure coordinates as listed in FIGURE 9.

74. The computer readable storage medium of claim 73, further comprising the identity of the atoms for the majority of the three-dimensional structure coordinates as listed in FIGURE 9.

75. The computer readable storage medium of claim 73, wherein the data comprise substantially all of the three-dimensional structure coordinates as listed in FIGURE 9.

1/76

FIGURE 1

SEQ ID NO: 1

ATGCGTCGACCCTTGGTGGCCGGTAACTGGAAAATGCACGGTACGCATT
CCAGTGTGGCCGAGTTGATCAAAGGCTTGCCTCAGCTGGCGTTGCCGAGCGGA
GTCGATGTGGCTGTGATGCCGCCTTGCTTGTTCATCAGCCAGGTCATCCAGGGC
CTGGCCGGCAAGGCGATCGATGTAGGTGCGCAGAACAGCGCCGTCGAGCCGAT
GCAAGGCGCGCTGACCGGTGAGACTGCTCCCAGTCAGTTGGCGGATGTCGGTT
GTAGCATGGTCCTTGTGGGCCACTCGGAGCGTCGCCTGATTCTCGGCGAGAGTG
ACGAGGTTGTGAGTCGCAAGTTTGCCGCGGCTCAGTCGTGCGGCCTGGTGCCGG
TGCTGTGTGTCGGGGAGACCCGGGCGGAGCGCGAGGCGGGCAAGACGCTGGA
GGTTGTGCGCAAGGCAGCTGGGAAGCGTGATCGACGAGTTGGGTGTTGGAGCTT
TTGCTCGCGCAGTCGTGGCTTACGAGCCGGTCTGGGCGATCGGGACCGGGTTGA
CCGCGTCCCCCGCGCAAGCCCAGGAAGTGCACGCGGCGATCCGCGCGCAACTG
GCAGCGGAAAATGCCGAGGTCGCAAAAGGTGTGCGACTCCTTTACGGCGGCAG
TGTAAGGCGGCAAGTGCAGCCGAGTTGTTGGCATGCCGGAATATCGATGGGG
GGCTGGTAGGTGGAGCCTCCCTCAATGCGGATGAGTTCGGCGCGATCTGTCGTG
CCGCGGGAAGCTGA

2/76

FIGURE 2

SEQ ID NO: 2

MRRPLVAGNWKMHGTHSSVABLIKGLRQLALPSGVDVAVMPPCLFISQVIQ
GLAGKAIDVGAQNSAVEPMQGALTGETAPSQLADVGC SMVLVGHSE RRLILGESD
EVVSRKFAAAQSCGLVPVLCVGETRAEREAGKTLEVVARQLG SVIDELGVGAFAR
AVVAYEPVWAIGTGLTASPAQAQEVHAAIRAQLAAENAEVAKGVRLLYGGSVKA
ASAAELFGMPDIDGGLVGGASLNADEFGAICRAAGS

3/76

FIGURE 3

SEQ ID NO: 3

ATGCGTCGACCCTTGGTGGCCGGTAACTGGAAAATGCACGGTACGCATT
CCAGTGTGGCCGAGTTGATCAAAGGCTTGCCTCAGCTGGCGTTGCCGAGCGGA
GTCGATGTAGCTGTGATGCCGCCTTGCTTGTTCATCAGCCAGGTCATCCAGGGC
CTGGCCGGCAAGGCGATCGATGTAGGTGCGCAGAACAGCGCCGTCGAGCCGAT
GCAAGGCGCGCTGACCGGTGAGACTGCTCCAGTCAGTTGGCGGATGTCGGTT
GTAGCATGGTCCTCGTGGGCCACTCGGAGCGTCGCCTGATTCTCGGCGAGAGTG
ACGAGGTTGTGAGTCGCAAGTTTGCCGCGGCTCAGTCGTGCGGCCTGGTGCCGG
TGCTGTGTGTCGGGGAGACCCGGGCGGAGCGCGAGGCGGGCAAGACGCTGGA
GGTTGTCGCAAGGCAGCTGGGAAGCGTGATCAACGAGTTGGGTGTTGGAGCTT
TTGCTCGCGCAGTCGTGGCTTACGAGCCGGTCTGGGCGATCGGGACCGGGTTGA
CCGCGTCCCCCGCGCAAGCCCAGGAAGTGACGCGGCGATCCGCGCGCAACTG
GCAGCGGAAAATGCCGAGGTCGAAAAGGTGTGCGACTCCTTTACGGCGGCAG
TGTAAGGCGGCAAGTGCAGCCGAGTTGTTTCGGCATGCCGGATATCGATGGGG
GGCTGGTAAGTGAGCCTCCCTCAATGCGGATGAGTTCGGCGCGATCTGTCGTG
CCGCGGGAAGCTGA

4/76

FIGURE 4

SEQ ID NO: 4

MRRPLVAGNWKMHGTHSSVAELIKGLRQLALPSGVDVAVMPPCLFISQVIQ
GLAGKAIDVGAQNSAVEPMQGALTGETAPSQLADVGC SMVLVGH SERRLILGESD
EVVSRKF AAAQSCGLVPVLCVGETRAERBAGKTLEVVARQLG SVINELGVGAFAR
AVVAYEPVWAIGTGLTASPAQAQEVHAAIRAQLAAENAEVAKGVRLLYGGSVKA
ASAAELFGMPDIDGGLVSGASLNADEFGAICRAAGS

576

FIGURE 5

SEQ ID NO: 5

Forward PCR Primer

GCGGCGGCATTAATATGAAAAATTGTGCATCGTCAG

SEQ ID NO: 6

Reverse PCR Primer

GCGCGGATCCTTATTAATTCAACCGTTCAATCACC

6/76

FIGURE 6

TABLE 1: Amino Acid and Nucleic Acid Properties

| | |
|--------------------------------------------------------------------------------------------------------------------------------------|------------------|
| Melting temperature (°C) of SEQ ID NO: 5 (forward PCR primer) | 62 |
| Restriction enzyme for SEQ ID NO: 5 (forward PCR primer) | AseI |
| Melting temperature (°C) of SEQ ID NO: 6 (reverse PCR primer) | 60 |
| Restriction enzyme for SEQ ID NO: 6 (reverse PCR primer) | BamHI |
| Number of nucleic acid residues in SEQ ID NO: 1 | 756 |
| Number of amino acid residues in SEQ ID NO: 2 | 251 |
| Number of different nucleic acid residues between SEQ ID NO: 1 and SEQ ID NO: 3 | 4 |
| Number of different amino acid residues between SEQ ID NO: 2 and SEQ ID NO: 4 | 2 |
| Calculated Molecular weight of SEQ ID NO: 2 polypeptide (kDa) | 27.9 |
| Calculated pI of SEQ ID NO: 2 polypeptide | 5.2 |
| Solubility of SEQ ID NO: 4 polypeptide, determined as described in EXAMPLE 2 (with the His tag at the N-terminus) | Approaching 100% |
| Amount of purified polypeptide having SEQ ID NO: 4, prepared and purified as described in EXAMPLE 7 (mg/L of culture) | 49.3 |
| Amount of purified selmet labeled polypeptide having SEQ ID NO: 4, prepared and purified as described in EXAMPLE 7 (mg/L of culture) | 123.2 |
| Amount of purified polypeptide having SEQ ID NO: 4 soluble in buffer, as described in EXAMPLE 7 (mg/ml of buffer) | 45.8 |
| Amount of purified selmet labeled polypeptide having SEQ ID NO: 4 soluble in buffer, as described in EXAMPLE 7 (mg/ml of buffer) | 91.6 |

FIGURE 7

TABLE 2: Bioinformatic Analyses

| | |
|---------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| Protein annotation and gene designation, if any | triosephosphate isomerase, <i>tpiA</i> |
| COG Category | carbohydrate transport and metabolism |
| COG ID Number | COG0149 |
| Is SEQ ID NO: 2 classified as an essential gene? | yes |
| Most closely related protein from PDB | Triosephosphate Isomerase, (1tre_B) |
| Source organism for closest PDB protein | <i>E. coli</i> |
| % Identity between SEQ ID NO: 2 and the closest protein from PDB | 52 |
| % Positives between SEQ ID NO: 2 and the closest protein from PDB | 67 |
| Number of Protein Hits in the VGDB | 12 |
| Number of Microorganisms having VGDB Hits | 12 |
| Microorganisms having VGDB Hits ¹ | [saur][paer][efae][bsub][spne][ctra] [ecoli][bbur][hinf][nmen][mgen][hpyl] |
| First predicted epitopic region of SEQ ID NO: 2: rank score, amino acid residue numbers; amino acid sequence | 1.304,105->129, SDEVVSRKF AAAQSCGLVPVLCVGE (SEQ ID NO: 10) |
| Second predicted epitopic region of SEQ ID NO: 2: rank score, amino acid residue numbers; amino acid sequence | 1.188,16->61, HSSVAELIKGLRQLALPSGVDVA VMPPCLFISQVIQGLAGK AIDVG (SEQ ID NO: 11) |
| Third predicted epitopic region of SEQ ID NO: 2: rank score, amino acid residue numbers; amino acid sequence | 1.184,139->175, TLEVVARQLG SVIDELGVG AFARAVVAYEPVWAIGTG (SEQ ID NO: 12) |

¹Organisms are abbreviated as follows: *ecoli* = *Eschericia coli*; *hpyl* = *Helicobacter pylori*; *paer* = *Pseudomonas aeruginosa*; *ctra* = *Chlaydia trachomatis*; *hinf* = *Haemophilus influenzae*; *nmen* = *Neisseria meningitidis*; *rpxx* = *Rickettsia prowazekii*; *bbur* = *Borrelia burgdorferi*; *bsub* = *Bacillus subtilis*; *staph* = *Staphylococcus aureus*; *spne* = *Streptococcus pneumoniae*; *mgen* = *Mycoplasma genitalium*; *efae* = *Enterococcus faecalis*.

FIGURE 8

TABLE 3: X-ray Structure Data

| | |
|----------------------------------------------|----------|
| Wavelength (Å) | 0.9796 |
| Resolution (Å) | 15 – 2.2 |
| Total Data | 202,194 |
| Unique Data | 28,770 |
| Redundancy | 7.5 |
| Overall Completeness* | 95.9 |
| Completeness (last shell)* | 90.3 |
| R _{merge} (overall) ^a | 8.8 |
| R _{merge} (last shell) ^a | 15.4 |
| <I/σ(I)> | 14.4 |

$$^a R_{\text{merge}} = \sum_{\text{unique reflections}} (\sum_{i=1}^N |I_i - \bar{I}|) / \sum_{\text{unique reflections}} (\sum_{i=1}^N I_i)$$

* last shell includes all reflections between 2.28 and 2.20 Å.

9/76

FIGURE 8 -2

TABLE 3: X-ray Structure Data Continued

| | | |
|------------------------------------------|------------------------------------------------------------|----------|
| <i>Final Model parameters</i> | Number of amino acid chains | 2 |
| | Number of protein atoms | 3576 |
| | Number of solvent molecules | 194 |
| | Resolution range (Å) | 15 – 2.2 |
| | R-factor ^b | 20.65 |
| | R _{free} | 25.03 |
| | Average main chain / side chain B-factor (Å ²) | 29.5 |
| | Average solvent B-factor (Å ²) | 35.4 |
| <i>RMS deviation from ideal geometry</i> | Covalent bond lengths (Å) | 0.010 |
| | Bond angles (°) | 1.288 |
| | Improper angles (°) | 0.842 |
| | Dihedral angles (°) | 22.587 |

^bR-factor = $\frac{\sum |F_{\text{obs}}| - |F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$

^cR_{free} is a cross-validation residual calculated using 5% of the native data which were chosen randomly and excluded from the refinement.

10/76

FIGURE 9 - 1

| | | | | | | | | | | | | |
|------|----|-----|-----|---|---|--------|--------|--------|------|-------|---|---|
| ATOM | 1 | CB | MET | A | 1 | 69.824 | 37.159 | 19.844 | 1.00 | 47.93 | A | C |
| ATOM | 2 | CG | MET | A | 1 | 70.963 | 36.609 | 20.697 | 1.00 | 51.83 | A | C |
| ATOM | 3 | SD | MET | A | 1 | 71.783 | 35.169 | 19.890 | 1.00 | 56.56 | A | S |
| ATOM | 4 | CE | MET | A | 1 | 73.200 | 36.024 | 18.993 | 1.00 | 54.75 | A | C |
| ATOM | 5 | C | MET | A | 1 | 69.421 | 35.932 | 17.696 | 1.00 | 42.10 | A | C |
| ATOM | 6 | O | MET | A | 1 | 68.625 | 36.084 | 16.759 | 1.00 | 43.43 | A | O |
| ATOM | 7 | N | MET | A | 1 | 69.697 | 38.417 | 17.710 | 1.00 | 45.33 | A | N |
| ATOM | 8 | CA | MET | A | 1 | 70.126 | 37.134 | 18.339 | 1.00 | 44.82 | A | C |
| ATOM | 9 | N | ARG | A | 2 | 69.734 | 34.738 | 18.186 | 1.00 | 37.50 | A | N |
| ATOM | 10 | CA | ARG | A | 2 | 69.122 | 33.518 | 17.688 | 1.00 | 33.20 | A | C |
| ATOM | 11 | CB | ARG | A | 2 | 70.096 | 32.342 | 17.808 | 1.00 | 31.53 | A | C |
| ATOM | 12 | CG | ARG | A | 2 | 71.365 | 32.478 | 16.972 | 1.00 | 30.51 | A | C |
| ATOM | 13 | CD | ARG | A | 2 | 72.265 | 31.262 | 17.156 | 1.00 | 25.81 | A | C |
| ATOM | 14 | NE | ARG | A | 2 | 72.686 | 31.138 | 18.543 | 1.00 | 26.87 | A | N |
| ATOM | 15 | CZ | ARG | A | 2 | 73.835 | 31.596 | 19.022 | 1.00 | 25.34 | A | C |
| ATOM | 16 | NH1 | ARG | A | 2 | 74.698 | 32.205 | 18.221 | 1.00 | 24.74 | A | N |
| ATOM | 17 | NH2 | ARG | A | 2 | 74.105 | 31.467 | 20.312 | 1.00 | 26.16 | A | N |
| ATOM | 18 | C | ARG | A | 2 | 67.885 | 33.245 | 18.536 | 1.00 | 30.48 | A | C |
| ATOM | 19 | O | ARG | A | 2 | 67.973 | 33.155 | 19.757 | 1.00 | 29.01 | A | O |
| ATOM | 20 | N | ARG | A | 3 | 66.737 | 33.119 | 17.883 | 1.00 | 28.55 | A | N |
| ATOM | 21 | CA | ARG | A | 3 | 65.476 | 32.861 | 18.569 | 1.00 | 29.41 | A | C |
| ATOM | 22 | CB | ARG | A | 3 | 64.332 | 33.036 | 17.575 | 1.00 | 30.70 | A | C |
| ATOM | 23 | CG | ARG | A | 3 | 62.936 | 33.054 | 18.153 | 1.00 | 31.57 | A | C |
| ATOM | 24 | CD | ARG | A | 3 | 61.944 | 33.346 | 17.030 | 1.00 | 32.95 | A | C |
| ATOM | 25 | NE | ARG | A | 3 | 60.551 | 33.310 | 17.463 | 1.00 | 35.33 | A | N |
| ATOM | 26 | CZ | ARG | A | 3 | 59.970 | 34.225 | 18.232 | 1.00 | 34.99 | A | C |
| ATOM | 27 | NH1 | ARG | A | 3 | 60.664 | 35.269 | 18.664 | 1.00 | 35.09 | A | N |
| ATOM | 28 | NH2 | ARG | A | 3 | 58.688 | 34.092 | 18.569 | 1.00 | 35.17 | A | N |
| ATOM | 29 | C | ARG | A | 3 | 65.490 | 31.438 | 19.131 | 1.00 | 28.70 | A | C |
| ATOM | 30 | O | ARG | A | 3 | 65.745 | 30.485 | 18.404 | 1.00 | 28.41 | A | O |
| ATOM | 31 | N | PRO | A | 4 | 65.234 | 31.278 | 20.440 | 1.00 | 28.03 | A | N |
| ATOM | 32 | CD | PRO | A | 4 | 64.972 | 32.307 | 21.460 | 1.00 | 28.75 | A | C |
| ATOM | 33 | CA | PRO | A | 4 | 65.234 | 29.940 | 21.044 | 1.00 | 27.75 | A | C |
| ATOM | 34 | CB | PRO | A | 4 | 64.912 | 30.220 | 22.512 | 1.00 | 28.22 | A | C |
| ATOM | 35 | CG | PRO | A | 4 | 65.433 | 31.615 | 22.721 | 1.00 | 27.20 | A | C |
| ATOM | 36 | C | PRO | A | 4 | 64.212 | 29.004 | 20.405 | 1.00 | 26.99 | A | C |
| ATOM | 37 | O | PRO | A | 4 | 63.154 | 29.443 | 19.951 | 1.00 | 26.58 | A | O |
| ATOM | 38 | N | LEU | A | 5 | 64.533 | 27.716 | 20.359 | 1.00 | 25.17 | A | N |
| ATOM | 39 | CA | LEU | A | 5 | 63.618 | 26.735 | 19.785 | 1.00 | 24.71 | A | C |
| ATOM | 40 | CB | LEU | A | 5 | 64.077 | 26.325 | 18.383 | 1.00 | 23.58 | A | C |
| ATOM | 41 | CG | LEU | A | 5 | 63.416 | 25.058 | 17.832 | 1.00 | 25.11 | A | C |
| ATOM | 42 | CD1 | LEU | A | 5 | 61.910 | 25.222 | 17.776 | 1.00 | 24.31 | A | C |
| ATOM | 43 | CD2 | LEU | A | 5 | 63.970 | 24.756 | 16.467 | 1.00 | 25.01 | A | C |
| ATOM | 44 | C | LEU | A | 5 | 63.489 | 25.494 | 20.669 | 1.00 | 23.84 | A | C |
| ATOM | 45 | O | LEU | A | 5 | 64.484 | 24.919 | 21.099 | 1.00 | 22.31 | A | O |
| ATOM | 46 | N | VAL | A | 6 | 62.250 | 25.106 | 20.949 | 1.00 | 22.72 | A | N |
| ATOM | 47 | CA | VAL | A | 6 | 61.980 | 23.926 | 21.762 | 1.00 | 22.86 | A | C |
| ATOM | 48 | CB | VAL | A | 6 | 61.255 | 24.284 | 23.081 | 1.00 | 23.69 | A | C |
| ATOM | 49 | CG1 | VAL | A | 6 | 60.942 | 23.015 | 23.855 | 1.00 | 24.56 | A | C |
| ATOM | 50 | CG2 | VAL | A | 6 | 62.127 | 25.203 | 23.925 | 1.00 | 21.51 | A | C |
| ATOM | 51 | C | VAL | A | 6 | 61.114 | 22.971 | 20.951 | 1.00 | 22.51 | A | C |
| ATOM | 52 | O | VAL | A | 6 | 59.976 | 23.280 | 20.617 | 1.00 | 21.30 | A | O |
| ATOM | 53 | N | ALA | A | 7 | 61.681 | 21.815 | 20.623 | 1.00 | 22.04 | A | N |
| ATOM | 54 | CA | ALA | A | 7 | 60.986 | 20.812 | 19.835 | 1.00 | 21.29 | A | C |
| ATOM | 55 | CB | ALA | A | 7 | 61.755 | 20.532 | 18.547 | 1.00 | 21.56 | A | C |
| ATOM | 56 | C | ALA | A | 7 | 60.829 | 19.537 | 20.639 | 1.00 | 20.53 | A | C |
| ATOM | 57 | O | ALA | A | 7 | 61.757 | 19.085 | 21.302 | 1.00 | 22.14 | A | O |
| ATOM | 58 | N | GLY | A | 8 | 59.638 | 18.968 | 20.573 | 1.00 | 20.59 | A | N |
| ATOM | 59 | CA | GLY | A | 8 | 59.360 | 17.749 | 21.296 | 1.00 | 21.98 | A | C |
| ATOM | 60 | C | GLY | A | 8 | 59.268 | 16.542 | 20.387 | 1.00 | 21.22 | A | C |
| ATOM | 61 | O | GLY | A | 8 | 58.391 | 16.458 | 19.524 | 1.00 | 20.48 | A | O |
| ATOM | 62 | N | ASN | A | 9 | 60.195 | 15.612 | 20.582 | 1.00 | 20.52 | A | N |
| ATOM | 63 | CA | ASN | A | 9 | 60.238 | 14.371 | 19.818 | 1.00 | 20.62 | A | C |

11/76

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 64 | CB | ASN | A | 9 | 61.688 | 13.905 | 19.664 | 1.00 | 21.79 | A | C |
| ATOM | 65 | CG | ASN | A | 9 | 61.814 | 12.604 | 18.898 | 1.00 | 22.57 | A | C |
| ATOM | 66 | OD1 | ASN | A | 9 | 60.860 | 12.131 | 18.267 | 1.00 | 23.21 | A | O |
| ATOM | 67 | ND2 | ASN | A | 9 | 63.007 | 12.021 | 18.938 | 1.00 | 20.11 | A | N |
| ATOM | 68 | C | ASN | A | 9 | 59.438 | 13.330 | 20.588 | 1.00 | 18.69 | A | C |
| ATOM | 69 | O | ASN | A | 9 | 59.893 | 12.804 | 21.599 | 1.00 | 18.72 | A | O |
| ATOM | 70 | N | TRP | A | 10 | 58.241 | 13.029 | 20.110 | 1.00 | 18.64 | A | N |
| ATOM | 71 | CA | TRP | A | 10 | 57.405 | 12.053 | 20.788 | 1.00 | 19.21 | A | C |
| ATOM | 72 | CB | TRP | A | 10 | 55.954 | 12.163 | 20.293 | 1.00 | 19.06 | A | C |
| ATOM | 73 | CG | TRP | A | 10 | 55.307 | 13.527 | 20.505 | 1.00 | 19.67 | A | C |
| ATOM | 74 | CD2 | TRP | A | 10 | 53.979 | 13.916 | 20.131 | 1.00 | 19.14 | A | C |
| ATOM | 75 | CE2 | TRP | A | 10 | 53.828 | 15.280 | 20.476 | 1.00 | 18.55 | A | C |
| ATOM | 76 | CE3 | TRP | A | 10 | 52.901 | 13.246 | 19.537 | 1.00 | 20.95 | A | C |
| ATOM | 77 | CD1 | TRP | A | 10 | 55.888 | 14.642 | 21.052 | 1.00 | 20.26 | A | C |
| ATOM | 78 | NE1 | TRP | A | 10 | 55.006 | 15.695 | 21.035 | 1.00 | 18.74 | A | N |
| ATOM | 79 | CE2 | TRP | A | 10 | 52.644 | 15.987 | 20.244 | 1.00 | 18.26 | A | C |
| ATOM | 80 | CZ3 | TRP | A | 10 | 51.721 | 13.949 | 19.307 | 1.00 | 22.89 | A | C |
| ATOM | 81 | CH2 | TRP | A | 10 | 51.605 | 15.311 | 19.661 | 1.00 | 19.41 | A | C |
| ATOM | 82 | C | TRP | A | 10 | 57.923 | 10.633 | 20.576 | 1.00 | 19.01 | A | C |
| ATOM | 83 | O | TRP | A | 10 | 57.626 | 9.740 | 21.360 | 1.00 | 21.00 | A | O |
| ATOM | 84 | N | LYS | A | 11 | 58.716 | 10.429 | 19.529 | 1.00 | 19.99 | A | N |
| ATOM | 85 | CA | LYS | A | 11 | 59.234 | 9.100 | 19.209 | 1.00 | 19.88 | A | C |
| ATOM | 86 | CB | LYS | A | 11 | 60.081 | 8.556 | 20.372 | 1.00 | 19.61 | A | C |
| ATOM | 87 | CG | LYS | A | 11 | 61.346 | 9.372 | 20.621 | 1.00 | 21.85 | A | C |
| ATOM | 88 | CD | LYS | A | 11 | 62.309 | 8.706 | 21.606 | 1.00 | 19.55 | A | C |
| ATOM | 89 | CE | LYS | A | 11 | 63.491 | 9.615 | 21.902 | 1.00 | 21.46 | A | C |
| ATOM | 90 | NZ | LYS | A | 11 | 64.357 | 9.126 | 23.015 | 1.00 | 21.50 | A | N |
| ATOM | 91 | C | LYS | A | 11 | 58.079 | 8.136 | 18.864 | 1.00 | 19.77 | A | C |
| ATOM | 92 | O | LYS | A | 11 | 57.011 | 8.578 | 18.442 | 1.00 | 19.57 | A | O |
| ATOM | 93 | N | MET | A | 12 | 58.293 | 6.836 | 19.047 | 1.00 | 19.49 | A | N |
| ATOM | 94 | CA | MET | A | 12 | 57.289 | 5.817 | 18.717 | 1.00 | 20.75 | A | C |
| ATOM | 95 | CB | MET | A | 12 | 57.992 | 4.469 | 18.499 | 1.00 | 17.93 | A | C |
| ATOM | 96 | CG | MET | A | 12 | 57.198 | 3.393 | 17.765 | 1.00 | 17.92 | A | C |
| ATOM | 97 | SD | MET | A | 12 | 58.141 | 1.848 | 17.540 | 1.00 | 9.74 | A | S |
| ATOM | 98 | CE | MET | A | 12 | 59.243 | 2.316 | 16.753 | 1.00 | 14.06 | A | C |
| ATOM | 99 | C | MET | A | 12 | 56.283 | 5.702 | 19.857 | 1.00 | 21.78 | A | C |
| ATOM | 100 | O | MET | A | 12 | 56.234 | 4.686 | 20.542 | 1.00 | 21.20 | A | O |
| ATOM | 101 | N | HIS | A | 13 | 55.482 | 6.748 | 20.051 | 1.00 | 22.25 | A | N |
| ATOM | 102 | CA | HIS | A | 13 | 54.501 | 6.772 | 21.132 | 1.00 | 22.79 | A | C |
| ATOM | 103 | CB | HIS | A | 13 | 55.025 | 7.623 | 22.297 | 1.00 | 21.95 | A | C |
| ATOM | 104 | CG | HIS | A | 13 | 56.214 | 7.042 | 22.999 | 1.00 | 22.25 | A | C |
| ATOM | 105 | CD2 | HIS | A | 13 | 57.533 | 7.343 | 22.918 | 1.00 | 20.66 | A | C |
| ATOM | 106 | ND1 | HIS | A | 13 | 56.111 | 6.020 | 23.918 | 1.00 | 19.93 | A | N |
| ATOM | 107 | CE1 | HIS | A | 13 | 57.313 | 5.718 | 24.374 | 1.00 | 22.47 | A | C |
| ATOM | 108 | NE2 | HIS | A | 13 | 58.194 | 6.506 | 23.782 | 1.00 | 21.78 | A | N |
| ATOM | 109 | C | HIS | A | 13 | 53.152 | 7.329 | 20.703 | 1.00 | 22.90 | A | C |
| ATOM | 110 | O | HIS | A | 13 | 53.059 | 8.095 | 19.748 | 1.00 | 25.02 | A | O |
| ATOM | 111 | N | GLY | A | 14 | 52.107 | 6.930 | 21.420 | 1.00 | 23.25 | A | N |
| ATOM | 112 | CA | GLY | A | 14 | 50.779 | 7.431 | 21.139 | 1.00 | 22.85 | A | C |
| ATOM | 113 | C | GLY | A | 14 | 49.759 | 6.512 | 20.499 | 1.00 | 23.50 | A | C |
| ATOM | 114 | O | GLY | A | 14 | 50.102 | 5.608 | 19.745 | 1.00 | 23.98 | A | O |
| ATOM | 115 | N | THR | A | 15 | 48.494 | 6.761 | 20.838 | 1.00 | 24.90 | A | N |
| ATOM | 116 | CA | THR | A | 15 | 47.331 | 6.053 | 20.301 | 1.00 | 27.25 | A | C |
| ATOM | 117 | CB | THR | A | 15 | 46.601 | 5.216 | 21.364 | 1.00 | 27.45 | A | C |
| ATOM | 118 | OG1 | THR | A | 15 | 46.225 | 6.067 | 22.457 | 1.00 | 28.53 | A | O |
| ATOM | 119 | CG2 | THR | A | 15 | 47.503 | 4.098 | 21.883 | 1.00 | 27.74 | A | C |
| ATOM | 120 | C | THR | A | 15 | 46.420 | 7.206 | 19.906 | 1.00 | 28.74 | A | C |
| ATOM | 121 | O | THR | A | 15 | 46.759 | 8.364 | 20.143 | 1.00 | 26.29 | A | O |
| ATOM | 122 | N | HIS | A | 16 | 45.265 | 6.916 | 19.324 | 1.00 | 30.03 | A | N |
| ATOM | 123 | CA | HIS | A | 16 | 44.387 | 8.005 | 18.924 | 1.00 | 32.19 | A | C |
| ATOM | 124 | CB | HIS | A | 16 | 43.164 | 7.481 | 18.157 | 1.00 | 34.83 | A | C |
| ATOM | 125 | CG | HIS | A | 16 | 42.518 | 8.524 | 17.301 | 1.00 | 38.07 | A | C |
| ATOM | 126 | CD2 | HIS | A | 16 | 41.488 | 9.370 | 17.549 | 1.00 | 38.27 | A | C |
| ATOM | 127 | ND1 | HIS | A | 16 | 43.014 | 8.875 | 16.061 | 1.00 | 38.87 | A | N |
| ATOM | 128 | CE1 | HIS | A | 16 | 42.318 | 9.895 | 15.586 | 1.00 | 39.88 | A | C |
| ATOM | 129 | NE2 | HIS | A | 16 | 41.388 | 10.216 | 16.469 | 1.00 | 38.72 | A | N |

FIGURE 9 - 2

SUBSTITUTE SHEET (RULE 26)

12/76

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 130 | C | HIS | A | 16 | 43.916 | 8.854 | 20.115 | 1.00 | 30.78 | A | C |
| ATOM | 131 | O | HIS | A | 16 | 43.917 | 10.087 | 20.052 | 1.00 | 29.88 | A | O |
| ATOM | 132 | N | SER | A | 17 | 43.537 | 8.199 | 21.205 | 1.00 | 30.04 | A | N |
| ATOM | 133 | CA | SER | A | 17 | 43.054 | 8.912 | 22.381 | 1.00 | 30.19 | A | C |
| ATOM | 134 | CB | SER | A | 17 | 42.272 | 7.962 | 23.281 | 1.00 | 31.56 | A | C |
| ATOM | 135 | OG | SER | A | 17 | 43.099 | 6.890 | 23.712 | 1.00 | 37.12 | A | O |
| ATOM | 136 | C | SER | A | 17 | 44.155 | 9.591 | 23.196 | 1.00 | 28.02 | A | C |
| ATOM | 137 | O | SER | A | 17 | 43.920 | 10.650 | 23.785 | 1.00 | 28.65 | A | O |
| ATOM | 138 | N | SER | A | 18 | 45.343 | 8.995 | 23.263 | 1.00 | 24.87 | A | N |
| ATOM | 139 | CA | SER | A | 18 | 46.412 | 9.638 | 24.022 | 1.00 | 24.92 | A | C |
| ATOM | 140 | CB | SER | A | 18 | 47.568 | 8.667 | 24.325 | 1.00 | 24.04 | A | C |
| ATOM | 141 | OG | SER | A | 18 | 48.271 | 8.292 | 23.160 | 1.00 | 26.18 | A | O |
| ATOM | 142 | C | SER | A | 18 | 46.912 | 10.853 | 23.234 | 1.00 | 23.58 | A | C |
| ATOM | 143 | O | SER | A | 18 | 47.250 | 11.885 | 23.822 | 1.00 | 22.83 | A | O |
| ATOM | 144 | N | VAL | A | 19 | 46.943 | 10.734 | 21.905 | 1.00 | 23.22 | A | N |
| ATOM | 145 | CA | VAL | A | 19 | 47.374 | 11.848 | 21.064 | 1.00 | 22.16 | A | C |
| ATOM | 146 | CB | VAL | A | 19 | 47.516 | 11.444 | 19.568 | 1.00 | 21.80 | A | C |
| ATOM | 147 | CG1 | VAL | A | 19 | 47.721 | 12.691 | 18.709 | 1.00 | 21.11 | A | C |
| ATOM | 148 | CG2 | VAL | A | 19 | 48.712 | 10.487 | 19.388 | 1.00 | 20.96 | A | C |
| ATOM | 149 | C | VAL | A | 19 | 46.336 | 12.963 | 21.189 | 1.00 | 22.14 | A | C |
| ATOM | 150 | O | VAL | A | 19 | 46.678 | 14.109 | 21.468 | 1.00 | 20.27 | A | O |
| ATOM | 151 | N | ALA | A | 20 | 45.067 | 12.613 | 20.996 | 1.00 | 24.06 | A | N |
| ATOM | 152 | CA | ALA | A | 20 | 43.975 | 13.579 | 21.098 | 1.00 | 24.10 | A | C |
| ATOM | 153 | CB | ALA | A | 20 | 42.638 | 12.862 | 21.036 | 1.00 | 23.57 | A | C |
| ATOM | 154 | C | ALA | A | 20 | 44.101 | 14.344 | 22.416 | 1.00 | 24.49 | A | C |
| ATOM | 155 | O | ALA | A | 20 | 43.978 | 15.570 | 22.452 | 1.00 | 23.70 | A | O |
| ATOM | 156 | N | GLU | A | 21 | 44.366 | 13.615 | 23.494 | 1.00 | 25.16 | A | N |
| ATOM | 157 | CA | GLU | A | 21 | 44.513 | 14.229 | 24.806 | 1.00 | 26.38 | A | C |
| ATOM | 158 | CB | GLU | A | 21 | 44.634 | 13.139 | 25.882 | 1.00 | 29.08 | A | C |
| ATOM | 159 | CG | GLU | A | 21 | 44.643 | 13.631 | 27.332 | 1.00 | 33.17 | A | C |
| ATOM | 160 | CD | GLU | A | 21 | 43.513 | 14.622 | 27.654 | 1.00 | 37.76 | A | C |
| ATOM | 161 | OE1 | GLU | A | 21 | 42.320 | 14.335 | 27.358 | 1.00 | 40.49 | A | O |
| ATOM | 162 | OE2 | GLU | A | 21 | 43.823 | 15.693 | 28.215 | 1.00 | 38.04 | A | O |
| ATOM | 163 | C | GLU | A | 21 | 45.718 | 15.176 | 24.842 | 1.00 | 25.82 | A | C |
| ATOM | 164 | O | GLU | A | 21 | 45.599 | 16.326 | 25.273 | 1.00 | 24.84 | A | O |
| ATOM | 165 | N | LEU | A | 22 | 46.872 | 14.708 | 24.377 | 1.00 | 24.48 | A | N |
| ATOM | 166 | CA | LEU | A | 22 | 48.068 | 15.547 | 24.366 | 1.00 | 24.06 | A | C |
| ATOM | 167 | CB | LEU | A | 22 | 49.233 | 14.779 | 23.747 | 1.00 | 24.45 | A | C |
| ATOM | 168 | CG | LEU | A | 22 | 50.577 | 15.481 | 23.531 | 1.00 | 24.83 | A | C |
| ATOM | 169 | CD1 | LEU | A | 22 | 51.042 | 16.173 | 24.809 | 1.00 | 23.46 | A | C |
| ATOM | 170 | CD2 | LEU | A | 22 | 51.600 | 14.429 | 23.068 | 1.00 | 23.04 | A | C |
| ATOM | 171 | C | LEU | A | 22 | 47.797 | 16.827 | 23.572 | 1.00 | 24.06 | A | C |
| ATOM | 172 | O | LEU | A | 22 | 48.239 | 17.917 | 23.951 | 1.00 | 23.44 | A | O |
| ATOM | 173 | N | ILE | A | 23 | 47.063 | 16.692 | 22.472 | 1.00 | 23.41 | A | N |
| ATOM | 174 | CA | ILE | A | 23 | 46.724 | 17.848 | 21.642 | 1.00 | 25.38 | A | C |
| ATOM | 175 | CB | ILE | A | 23 | 45.848 | 17.462 | 20.424 | 1.00 | 25.03 | A | C |
| ATOM | 176 | CG2 | ILE | A | 23 | 45.422 | 18.723 | 19.669 | 1.00 | 23.51 | A | C |
| ATOM | 177 | CG1 | ILE | A | 23 | 46.611 | 16.513 | 19.503 | 1.00 | 23.69 | A | C |
| ATOM | 178 | CD1 | ILE | A | 23 | 45.765 | 15.973 | 18.387 | 1.00 | 23.55 | A | C |
| ATOM | 179 | C | ILE | A | 23 | 45.918 | 18.841 | 22.464 | 1.00 | 26.93 | A | C |
| ATOM | 180 | O | ILE | A | 23 | 46.182 | 20.044 | 22.445 | 1.00 | 26.59 | A | O |
| ATOM | 181 | N | LYS | A | 24 | 44.922 | 18.324 | 23.179 | 1.00 | 28.05 | A | N |
| ATOM | 182 | CA | LYS | A | 24 | 44.074 | 19.173 | 23.996 | 1.00 | 28.73 | A | C |
| ATOM | 183 | CB | LYS | A | 24 | 42.985 | 18.356 | 24.696 | 1.00 | 30.66 | A | C |
| ATOM | 184 | CG | LYS | A | 24 | 41.815 | 19.229 | 25.111 | 1.00 | 34.94 | A | C |
| ATOM | 185 | CD | LYS | A | 24 | 40.601 | 18.433 | 25.579 | 1.00 | 38.91 | A | C |
| ATOM | 186 | CE | LYS | A | 24 | 39.287 | 19.094 | 25.099 | 1.00 | 40.17 | A | C |
| ATOM | 187 | NZ | LYS | A | 24 | 39.076 | 20.525 | 25.510 | 1.00 | 39.57 | A | N |
| ATOM | 188 | C | LYS | A | 24 | 44.887 | 19.936 | 25.023 | 1.00 | 27.19 | A | C |
| ATOM | 189 | O | LYS | A | 24 | 44.660 | 21.129 | 25.236 | 1.00 | 27.16 | A | O |
| ATOM | 190 | N | GLY | A | 25 | 45.837 | 19.253 | 25.654 | 1.00 | 26.00 | A | N |
| ATOM | 191 | CA | GLY | A | 25 | 46.670 | 19.902 | 26.645 | 1.00 | 24.59 | A | C |
| ATOM | 192 | C | GLY | A | 25 | 47.560 | 20.974 | 26.038 | 1.00 | 25.36 | A | C |
| ATOM | 193 | O | GLY | A | 25 | 47.803 | 22.003 | 26.665 | 1.00 | 26.41 | A | O |
| ATOM | 194 | N | LEU | A | 26 | 48.050 | 20.738 | 24.820 | 1.00 | 24.53 | A | N |
| ATOM | 195 | CA | LEU | A | 26 | 48.915 | 21.703 | 24.142 | 1.00 | 24.27 | A | C |

FIGURE 9 - 3

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 196 | CB | LEU | A | 26 | 49.518 | 21.086 | 22.870 | 1.00 | 23.24 | A | C |
| ATOM | 197 | CG | LEU | A | 26 | 50.587 | 19.999 | 23.059 | 1.00 | 23.26 | A | C |
| ATOM | 198 | CD1 | LEU | A | 26 | 50.912 | 19.349 | 21.714 | 1.00 | 21.18 | A | C |
| ATOM | 199 | CD2 | LEU | A | 26 | 51.840 | 20.602 | 23.689 | 1.00 | 22.07 | A | C |
| ATOM | 200 | C | LEU | A | 26 | 48.169 | 22.998 | 23.793 | 1.00 | 23.31 | A | C |
| ATOM | 201 | O | LEU | A | 26 | 48.696 | 24.087 | 23.979 | 1.00 | 22.24 | A | O |
| ATOM | 202 | N | ARG | A | 27 | 46.945 | 22.873 | 23.290 | 1.00 | 23.81 | A | N |
| ATOM | 203 | CA | ARG | A | 27 | 46.140 | 24.037 | 22.932 | 1.00 | 23.88 | A | C |
| ATOM | 204 | CB | ARG | A | 27 | 44.762 | 23.611 | 22.427 | 1.00 | 23.32 | A | C |
| ATOM | 205 | CG | ARG | A | 27 | 44.741 | 22.851 | 21.128 | 1.00 | 26.57 | A | C |
| ATOM | 206 | CD | ARG | A | 27 | 43.313 | 22.489 | 20.770 | 1.00 | 27.06 | A | C |
| ATOM | 207 | NE | ARG | A | 27 | 42.504 | 23.674 | 20.505 | 1.00 | 30.39 | A | N |
| ATOM | 208 | CZ | ARG | A | 27 | 42.544 | 24.374 | 19.375 | 1.00 | 32.04 | A | C |
| ATOM | 209 | NH1 | ARG | A | 27 | 43.356 | 24.017 | 18.387 | 1.00 | 30.90 | A | N |
| ATOM | 210 | NH2 | ARG | A | 27 | 41.752 | 25.426 | 19.224 | 1.00 | 31.37 | A | N |
| ATOM | 211 | C | ARG | A | 27 | 45.930 | 24.922 | 24.155 | 1.00 | 24.82 | A | C |
| ATOM | 212 | O | ARG | A | 27 | 45.895 | 26.145 | 24.052 | 1.00 | 24.98 | A | O |
| ATOM | 213 | N | GLN | A | 28 | 45.782 | 24.286 | 25.311 | 1.00 | 26.47 | A | N |
| ATOM | 214 | CA | GLN | A | 28 | 45.542 | 24.994 | 26.566 | 1.00 | 28.33 | A | C |
| ATOM | 215 | CB | GLN | A | 28 | 45.071 | 24.008 | 27.637 | 1.00 | 30.08 | A | C |
| ATOM | 216 | CG | GLN | A | 28 | 43.809 | 23.238 | 27.299 | 1.00 | 33.15 | A | C |
| ATOM | 217 | CD | GLN | A | 28 | 42.562 | 24.099 | 27.344 | 1.00 | 34.27 | A | C |
| ATOM | 218 | OE1 | GLN | A | 28 | 42.286 | 24.754 | 28.355 | 1.00 | 34.67 | A | O |
| ATOM | 219 | NE2 | GLN | A | 28 | 41.797 | 24.100 | 26.250 | 1.00 | 32.10 | A | N |
| ATOM | 220 | C | GLN | A | 28 | 46.759 | 25.740 | 27.109 | 1.00 | 28.85 | A | C |
| ATOM | 221 | O | GLN | A | 28 | 46.622 | 26.636 | 27.944 | 1.00 | 29.20 | A | O |
| ATOM | 222 | N | LEU | A | 29 | 47.942 | 25.369 | 26.637 | 1.00 | 28.41 | A | N |
| ATOM | 223 | CA | LEU | A | 29 | 49.173 | 25.975 | 27.116 | 1.00 | 27.29 | A | C |
| ATOM | 224 | CB | LEU | A | 29 | 50.389 | 25.267 | 26.503 | 1.00 | 27.49 | A | C |
| ATOM | 225 | CG | LEU | A | 29 | 50.579 | 23.795 | 26.863 | 1.00 | 29.30 | A | C |
| ATOM | 226 | CD1 | LEU | A | 29 | 51.881 | 23.297 | 26.249 | 1.00 | 28.82 | A | C |
| ATOM | 227 | CD2 | LEU | A | 29 | 50.608 | 23.627 | 28.389 | 1.00 | 29.57 | A | C |
| ATOM | 228 | C | LEU | A | 29 | 49.322 | 27.464 | 26.891 | 1.00 | 26.33 | A | C |
| ATOM | 229 | O | LEU | A | 29 | 48.897 | 28.006 | 25.874 | 1.00 | 25.97 | A | O |
| ATOM | 230 | N | ALA | A | 30 | 49.945 | 28.114 | 27.864 | 1.00 | 26.07 | A | N |
| ATOM | 231 | CA | ALA | A | 30 | 50.219 | 29.535 | 27.791 | 1.00 | 26.78 | A | C |
| ATOM | 232 | CB | ALA | A | 30 | 50.103 | 30.168 | 29.173 | 1.00 | 27.51 | A | C |
| ATOM | 233 | C | ALA | A | 30 | 51.651 | 29.631 | 27.284 | 1.00 | 26.63 | A | C |
| ATOM | 234 | O | ALA | A | 30 | 52.591 | 29.758 | 28.067 | 1.00 | 28.51 | A | O |
| ATOM | 235 | N | LEU | A | 31 | 51.816 | 29.550 | 25.970 | 1.00 | 26.79 | A | N |
| ATOM | 236 | CA | LEU | A | 31 | 53.145 | 29.618 | 25.368 | 1.00 | 27.63 | A | C |
| ATOM | 237 | CB | LEU | A | 31 | 53.076 | 29.304 | 23.869 | 1.00 | 25.11 | A | C |
| ATOM | 238 | CG | LEU | A | 31 | 52.600 | 27.885 | 23.563 | 1.00 | 26.13 | A | C |
| ATOM | 239 | CD1 | LEU | A | 31 | 52.642 | 27.646 | 22.067 | 1.00 | 22.59 | A | C |
| ATOM | 240 | CD2 | LEU | A | 31 | 53.486 | 26.877 | 24.324 | 1.00 | 24.79 | A | C |
| ATOM | 241 | C | LEU | A | 31 | 53.792 | 30.978 | 25.573 | 1.00 | 26.83 | A | C |
| ATOM | 242 | O | LEU | A | 31 | 53.115 | 31.999 | 25.611 | 1.00 | 27.18 | A | O |
| ATOM | 243 | N | PRO | A | 32 | 55.121 | 31.002 | 25.702 | 1.00 | 26.43 | A | N |
| ATOM | 244 | CD | PRO | A | 32 | 56.043 | 29.856 | 25.780 | 1.00 | 28.08 | A | C |
| ATOM | 245 | CA | PRO | A | 32 | 55.840 | 32.255 | 25.906 | 1.00 | 27.72 | A | C |
| ATOM | 246 | CB | PRO | A | 32 | 57.129 | 31.792 | 26.564 | 1.00 | 28.06 | A | C |
| ATOM | 247 | CG | PRO | A | 32 | 57.414 | 30.527 | 25.808 | 1.00 | 26.92 | A | C |
| ATOM | 248 | C | PRO | A | 32 | 56.101 | 32.940 | 24.588 | 1.00 | 28.93 | A | C |
| ATOM | 249 | O | PRO | A | 32 | 56.120 | 32.293 | 23.540 | 1.00 | 29.89 | A | O |
| ATOM | 250 | N | SER | A | 33 | 56.303 | 34.250 | 24.649 | 1.00 | 28.33 | A | N |
| ATOM | 251 | CA | SER | A | 33 | 56.592 | 35.029 | 23.463 | 1.00 | 30.67 | A | C |
| ATOM | 252 | CB | SER | A | 33 | 56.178 | 36.482 | 23.691 | 1.00 | 32.61 | A | C |
| ATOM | 253 | OG | SER | A | 33 | 56.600 | 36.904 | 24.975 | 1.00 | 36.43 | A | O |
| ATOM | 254 | C | SER | A | 33 | 58.089 | 34.954 | 23.185 | 1.00 | 29.53 | A | C |
| ATOM | 255 | O | SER | A | 33 | 58.877 | 34.703 | 24.089 | 1.00 | 31.31 | A | O |
| ATOM | 256 | N | GLY | A | 34 | 58.476 | 35.161 | 21.932 | 1.00 | 28.73 | A | N |
| ATOM | 257 | CA | GLY | A | 34 | 59.884 | 35.120 | 21.575 | 1.00 | 28.18 | A | C |
| ATOM | 258 | C | GLY | A | 34 | 60.517 | 33.740 | 21.574 | 1.00 | 27.91 | A | C |
| ATOM | 259 | O | GLY | A | 34 | 61.742 | 33.622 | 21.595 | 1.00 | 28.24 | A | O |
| ATOM | 260 | N | VAL | A | 35 | 59.693 | 32.698 | 21.530 | 1.00 | 28.52 | A | N |
| ATOM | 261 | CA | VAL | A | 35 | 60.192 | 31.325 | 21.529 | 1.00 | 27.99 | A | C |

FIGURE 9 - 4

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 262 | CB | VAL | A | 35 | 60.046 | 30.675 | 22.931 | 1.00 | 29.48 | A | C |
| ATOM | 263 | CG1 | VAL | A | 35 | 60.341 | 29.177 | 22.853 | 1.00 | 28.79 | A | C |
| ATOM | 264 | CG2 | VAL | A | 35 | 60.982 | 31.349 | 23.915 | 1.00 | 28.47 | A | C |
| ATOM | 265 | C | VAL | A | 35 | 59.461 | 30.443 | 20.523 | 1.00 | 27.27 | A | C |
| ATOM | 266 | O | VAL | A | 35 | 58.245 | 30.256 | 20.612 | 1.00 | 26.49 | A | O |
| ATOM | 267 | N | ASP | A | 36 | 60.195 | 29.899 | 19.561 | 1.00 | 26.91 | A | N |
| ATOM | 268 | CA | ASP | A | 36 | 59.560 | 29.021 | 18.591 | 1.00 | 26.67 | A | C |
| ATOM | 269 | CB | ASP | A | 36 | 60.441 | 28.819 | 17.355 | 1.00 | 29.43 | A | C |
| ATOM | 270 | CG | ASP | A | 36 | 60.608 | 30.093 | 16.530 | 1.00 | 31.48 | A | C |
| ATOM | 271 | OD1 | ASP | A | 36 | 59.601 | 30.792 | 16.278 | 1.00 | 32.70 | A | O |
| ATOM | 272 | OD2 | ASP | A | 36 | 61.751 | 30.387 | 16.117 | 1.00 | 32.67 | A | O |
| ATOM | 273 | C | ASP | A | 36 | 59.308 | 27.676 | 19.271 | 1.00 | 25.45 | A | C |
| ATOM | 274 | O | ASP | A | 36 | 60.078 | 27.243 | 20.132 | 1.00 | 24.80 | A | O |
| ATOM | 275 | N | VAL | A | 37 | 58.221 | 27.027 | 18.889 | 1.00 | 22.78 | A | N |
| ATOM | 276 | CA | VAL | A | 37 | 57.876 | 25.742 | 19.462 | 1.00 | 23.34 | A | C |
| ATOM | 277 | CB | VAL | A | 37 | 56.697 | 25.859 | 20.450 | 1.00 | 22.08 | A | C |
| ATOM | 278 | CG1 | VAL | A | 37 | 56.411 | 24.495 | 21.081 | 1.00 | 22.51 | A | C |
| ATOM | 279 | CG2 | VAL | A | 37 | 57.009 | 26.887 | 21.510 | 1.00 | 19.02 | A | C |
| ATOM | 280 | C | VAL | A | 37 | 57.486 | 24.782 | 18.350 | 1.00 | 23.64 | A | C |
| ATOM | 281 | O | VAL | A | 37 | 56.820 | 25.167 | 17.386 | 1.00 | 22.03 | A | O |
| ATOM | 282 | N | ALA | A | 38 | 57.910 | 23.530 | 18.488 | 1.00 | 23.04 | A | N |
| ATOM | 283 | CA | ALA | A | 38 | 57.593 | 22.516 | 17.494 | 1.00 | 23.13 | A | C |
| ATOM | 284 | CB | ALA | A | 38 | 58.691 | 22.462 | 16.431 | 1.00 | 21.25 | A | C |
| ATOM | 285 | C | ALA | A | 38 | 57.444 | 21.158 | 18.169 | 1.00 | 23.01 | A | C |
| ATOM | 286 | O | ALA | A | 38 | 57.976 | 20.930 | 19.255 | 1.00 | 21.43 | A | O |
| ATOM | 287 | N | VAL | A | 39 | 56.685 | 20.274 | 17.530 | 1.00 | 22.43 | A | N |
| ATOM | 288 | CA | VAL | A | 39 | 56.481 | 18.924 | 18.029 | 1.00 | 20.43 | A | C |
| ATOM | 289 | CB | VAL | A | 39 | 55.090 | 18.749 | 18.699 | 1.00 | 21.20 | A | C |
| ATOM | 290 | CG1 | VAL | A | 39 | 55.031 | 19.541 | 20.008 | 1.00 | 19.08 | A | C |
| ATOM | 291 | CG2 | VAL | A | 39 | 53.992 | 19.190 | 17.740 | 1.00 | 17.28 | A | C |
| ATOM | 292 | C | VAL | A | 39 | 56.569 | 17.985 | 16.835 | 1.00 | 21.10 | A | C |
| ATOM | 293 | O | VAL | A | 39 | 56.112 | 18.305 | 15.729 | 1.00 | 19.81 | A | O |
| ATOM | 294 | N | MET | A | 40 | 57.168 | 16.825 | 17.055 | 1.00 | 21.30 | A | N |
| ATOM | 295 | CA | MET | A | 40 | 57.305 | 15.852 | 15.990 | 1.00 | 22.14 | A | C |
| ATOM | 296 | CB | MET | A | 40 | 58.780 | 15.576 | 15.720 | 1.00 | 20.47 | A | C |
| ATOM | 297 | CG | MET | A | 40 | 59.530 | 16.775 | 15.163 | 1.00 | 22.33 | A | C |
| ATOM | 298 | SD | MET | A | 40 | 59.869 | 18.035 | 16.366 | 1.00 | 19.84 | A | S |
| ATOM | 299 | CE | MET | A | 40 | 60.409 | 19.333 | 15.287 | 1.00 | 22.36 | A | C |
| ATOM | 300 | C | MET | A | 40 | 56.564 | 14.571 | 16.365 | 1.00 | 22.82 | A | C |
| ATOM | 301 | O | MET | A | 40 | 57.110 | 13.684 | 17.028 | 1.00 | 20.11 | A | O |
| ATOM | 302 | N | PRO | A | 41 | 55.293 | 14.472 | 15.957 | 1.00 | 22.74 | A | N |
| ATOM | 303 | CD | PRO | A | 41 | 54.431 | 15.536 | 15.413 | 1.00 | 23.30 | A | C |
| ATOM | 304 | CA | PRO | A | 41 | 54.508 | 13.282 | 16.274 | 1.00 | 23.61 | A | C |
| ATOM | 305 | CB | PRO | A | 41 | 53.075 | 13.800 | 16.231 | 1.00 | 23.11 | A | C |
| ATOM | 306 | CG | PRO | A | 41 | 53.128 | 14.813 | 15.162 | 1.00 | 22.29 | A | C |
| ATOM | 307 | C | PRO | A | 41 | 54.741 | 12.133 | 15.307 | 1.00 | 23.69 | A | C |
| ATOM | 308 | O | PRO | A | 41 | 55.162 | 12.336 | 14.169 | 1.00 | 22.83 | A | O |
| ATOM | 309 | N | PRO | A | 42 | 54.489 | 10.900 | 15.765 | 1.00 | 24.15 | A | N |
| ATOM | 310 | CD | PRO | A | 42 | 54.112 | 10.480 | 17.126 | 1.00 | 23.29 | A | C |
| ATOM | 311 | CA | PRO | A | 42 | 54.677 | 9.740 | 14.902 | 1.00 | 22.61 | A | C |
| ATOM | 312 | CB | PRO | A | 42 | 54.095 | 8.607 | 15.730 | 1.00 | 23.16 | A | C |
| ATOM | 313 | CG | PRO | A | 42 | 54.458 | 9.005 | 17.109 | 1.00 | 23.14 | A | C |
| ATOM | 314 | C | PRO | A | 42 | 53.900 | 9.990 | 13.629 | 1.00 | 23.47 | A | C |
| ATOM | 315 | O | PRO | A | 42 | 52.844 | 10.619 | 13.640 | 1.00 | 23.31 | A | O |
| ATOM | 316 | N | CYS | A | 43 | 54.433 | 9.494 | 12.530 | 1.00 | 24.33 | A | N |
| ATOM | 317 | CA | CYS | A | 43 | 53.820 | 9.670 | 11.237 | 1.00 | 27.11 | A | C |
| ATOM | 318 | CB | CYS | A | 43 | 54.557 | 8.806 | 10.222 | 1.00 | 30.83 | A | C |
| ATOM | 319 | SG | CYS | A | 43 | 53.822 | 8.913 | 8.615 | 1.00 | 42.62 | A | S |
| ATOM | 320 | C | CYS | A | 43 | 52.308 | 9.430 | 11.117 | 1.00 | 25.98 | A | C |
| ATOM | 321 | O | CYS | A | 43 | 51.608 | 10.235 | 10.508 | 1.00 | 25.11 | A | O |
| ATOM | 322 | N | LEU | A | 44 | 51.777 | 8.360 | 11.697 | 1.00 | 25.14 | A | N |
| ATOM | 323 | CA | LEU | A | 44 | 50.346 | 8.131 | 11.521 | 1.00 | 27.46 | A | C |
| ATOM | 324 | CB | LEU | A | 44 | 49.986 | 6.660 | 11.792 | 1.00 | 29.68 | A | C |
| ATOM | 325 | CG | LEU | A | 44 | 49.818 | 6.186 | 13.221 | 1.00 | 30.68 | A | C |
| ATOM | 326 | CD1 | LEU | A | 44 | 49.676 | 4.677 | 13.244 | 1.00 | 31.49 | A | C |
| ATOM | 327 | CD2 | LEU | A | 44 | 51.010 | 6.622 | 14.020 | 1.00 | 34.87 | A | C |

FIGURE 9 - 5

SUBSTITUTE SHEET (RULE 26)

15/76

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 328 | C | LEU | A | 44 | 49.401 | 9.059 | 12.287 | 1.00 | 24.91 | A | C |
| ATOM | 329 | O | LEU | A | 44 | 48.199 | 8.997 | 12.084 | 1.00 | 23.50 | A | O |
| ATOM | 330 | N | PHE | A | 45 | 49.943 | 9.924 | 13.144 | 1.00 | 25.04 | A | N |
| ATOM | 331 | CA | PHE | A | 45 | 49.125 | 10.880 | 13.897 | 1.00 | 24.67 | A | C |
| ATOM | 332 | CB | PHE | A | 45 | 49.378 | 10.756 | 15.404 | 1.00 | 24.26 | A | C |
| ATOM | 333 | CG | PHE | A | 45 | 49.107 | 9.390 | 15.958 | 1.00 | 24.13 | A | C |
| ATOM | 334 | CD1 | PHE | A | 45 | 50.132 | 8.649 | 16.539 | 1.00 | 23.46 | A | C |
| ATOM | 335 | CD2 | PHE | A | 45 | 47.836 | 8.832 | 15.878 | 1.00 | 23.92 | A | C |
| ATOM | 336 | CE1 | PHE | A | 45 | 49.902 | 7.372 | 17.029 | 1.00 | 21.97 | A | C |
| ATOM | 337 | CE2 | PHE | A | 45 | 47.591 | 7.548 | 16.365 | 1.00 | 24.39 | A | C |
| ATOM | 338 | CZ | PHE | A | 45 | 48.626 | 6.817 | 16.940 | 1.00 | 26.49 | A | C |
| ATOM | 339 | C | PHE | A | 45 | 49.429 | 12.322 | 13.483 | 1.00 | 24.12 | A | C |
| ATOM | 340 | O | PHE | A | 45 | 48.890 | 13.263 | 14.067 | 1.00 | 26.74 | A | O |
| ATOM | 341 | N | ILE | A | 46 | 50.290 | 12.497 | 12.485 | 1.00 | 23.68 | A | N |
| ATOM | 342 | CA | ILE | A | 46 | 50.671 | 13.835 | 12.038 | 1.00 | 24.20 | A | C |
| ATOM | 343 | CB | ILE | A | 46 | 51.741 | 13.774 | 10.913 | 1.00 | 23.87 | A | C |
| ATOM | 344 | CG2 | ILE | A | 46 | 51.928 | 15.154 | 10.283 | 1.00 | 23.95 | A | C |
| ATOM | 345 | CG1 | ILE | A | 46 | 53.077 | 13.294 | 11.500 | 1.00 | 25.33 | A | C |
| ATOM | 346 | CD1 | ILE | A | 46 | 54.197 | 13.119 | 10.493 | 1.00 | 24.22 | A | C |
| ATOM | 347 | C | ILE | A | 46 | 49.487 | 14.680 | 11.581 | 1.00 | 25.62 | A | C |
| ATOM | 348 | O | ILE | A | 46 | 49.361 | 15.838 | 11.982 | 1.00 | 26.17 | A | O |
| ATOM | 349 | N | SER | A | 47 | 48.619 | 14.119 | 10.746 | 1.00 | 24.63 | A | N |
| ATOM | 350 | CA | SER | A | 47 | 47.458 | 14.875 | 10.291 | 1.00 | 26.08 | A | C |
| ATOM | 351 | CB | SER | A | 47 | 46.693 | 14.100 | 9.219 | 1.00 | 27.67 | A | C |
| ATOM | 352 | OG | SER | A | 47 | 47.343 | 14.270 | 7.967 | 1.00 | 33.06 | A | O |
| ATOM | 353 | C | SER | A | 47 | 46.541 | 15.226 | 11.454 | 1.00 | 25.19 | A | C |
| ATOM | 354 | O | SER | A | 47 | 46.039 | 16.343 | 11.533 | 1.00 | 26.32 | A | O |
| ATOM | 355 | N | GLN | A | 48 | 46.337 | 14.280 | 12.364 | 1.00 | 25.96 | A | N |
| ATOM | 356 | CA | GLN | A | 48 | 45.501 | 14.524 | 13.539 | 1.00 | 26.76 | A | C |
| ATOM | 357 | CB | GLN | A | 48 | 45.452 | 13.274 | 14.425 | 1.00 | 27.53 | A | C |
| ATOM | 358 | CG | GLN | A | 48 | 44.551 | 13.400 | 15.659 | 1.00 | 28.85 | A | C |
| ATOM | 359 | CD | GLN | A | 48 | 44.586 | 12.149 | 16.517 | 1.00 | 29.71 | A | C |
| ATOM | 360 | OE1 | GLN | A | 48 | 45.041 | 11.097 | 16.067 | 1.00 | 29.57 | A | O |
| ATOM | 361 | NE2 | GLN | A | 48 | 44.100 | 12.252 | 17.752 | 1.00 | 30.07 | A | N |
| ATOM | 362 | C | GLN | A | 48 | 46.058 | 15.700 | 14.355 | 1.00 | 26.95 | A | C |
| ATOM | 363 | O | GLN | A | 48 | 45.306 | 16.569 | 14.816 | 1.00 | 26.69 | A | O |
| ATOM | 364 | N | VAL | A | 49 | 47.379 | 15.725 | 14.528 | 1.00 | 24.70 | A | N |
| ATOM | 365 | CA | VAL | A | 49 | 48.024 | 16.785 | 15.296 | 1.00 | 22.94 | A | C |
| ATOM | 366 | CB | VAL | A | 49 | 49.497 | 16.424 | 15.590 | 1.00 | 22.84 | A | C |
| ATOM | 367 | CG1 | VAL | A | 49 | 50.219 | 17.602 | 16.219 | 1.00 | 21.67 | A | C |
| ATOM | 368 | CG2 | VAL | A | 49 | 49.545 | 15.229 | 16.536 | 1.00 | 22.25 | A | C |
| ATOM | 369 | C | VAL | A | 49 | 47.938 | 18.149 | 14.611 | 1.00 | 22.11 | A | C |
| ATOM | 370 | O | VAL | A | 49 | 47.638 | 19.146 | 15.262 | 1.00 | 21.72 | A | O |
| ATOM | 371 | N | ILE | A | 50 | 48.190 | 18.191 | 13.305 | 1.00 | 21.90 | A | N |
| ATOM | 372 | CA | ILE | A | 50 | 48.124 | 19.446 | 12.550 | 1.00 | 23.35 | A | C |
| ATOM | 373 | CB | ILE | A | 50 | 48.538 | 19.242 | 11.072 | 1.00 | 22.17 | A | C |
| ATOM | 374 | CG2 | ILE | A | 50 | 48.172 | 20.464 | 10.250 | 1.00 | 20.62 | A | C |
| ATOM | 375 | CG1 | ILE | A | 50 | 50.043 | 18.953 | 10.990 | 1.00 | 23.59 | A | C |
| ATOM | 376 | CD1 | ILE | A | 50 | 50.559 | 18.699 | 9.576 | 1.00 | 23.85 | A | C |
| ATOM | 377 | C | ILE | A | 50 | 46.706 | 20.024 | 12.593 | 1.00 | 25.09 | A | C |
| ATOM | 378 | O | ILE | A | 50 | 46.517 | 21.206 | 12.872 | 1.00 | 25.14 | A | O |
| ATOM | 379 | N | GLN | A | 51 | 45.713 | 19.187 | 12.308 | 1.00 | 26.88 | A | N |
| ATOM | 380 | CA | GLN | A | 51 | 44.322 | 19.630 | 12.343 | 1.00 | 29.22 | A | C |
| ATOM | 381 | CB | GLN | A | 51 | 43.386 | 18.488 | 11.934 | 1.00 | 32.23 | A | C |
| ATOM | 382 | CG | GLN | A | 51 | 43.185 | 18.336 | 10.430 | 1.00 | 38.55 | A | C |
| ATOM | 383 | CD | GLN | A | 51 | 42.375 | 17.089 | 10.060 | 1.00 | 43.90 | A | C |
| ATOM | 384 | OE1 | GLN | A | 51 | 41.489 | 16.658 | 10.813 | 1.00 | 44.08 | A | O |
| ATOM | 385 | NE2 | GLN | A | 51 | 42.668 | 16.513 | 8.884 | 1.00 | 45.38 | A | N |
| ATOM | 386 | C | GLN | A | 51 | 43.965 | 20.106 | 13.750 | 1.00 | 28.03 | A | C |
| ATOM | 387 | O | GLN | A | 51 | 43.331 | 21.142 | 13.915 | 1.00 | 28.31 | A | O |
| ATOM | 388 | N | GLY | A | 52 | 44.401 | 19.350 | 14.755 | 1.00 | 27.06 | A | N |
| ATOM | 389 | CA | GLY | A | 52 | 44.110 | 19.682 | 16.141 | 1.00 | 24.88 | A | C |
| ATOM | 390 | C | GLY | A | 52 | 44.729 | 20.948 | 16.718 | 1.00 | 25.06 | A | C |
| ATOM | 391 | O | GLY | A | 52 | 44.094 | 21.625 | 17.528 | 1.00 | 23.33 | A | O |
| ATOM | 392 | N | LEU | A | 53 | 45.958 | 21.268 | 16.317 | 1.00 | 24.55 | A | N |
| ATOM | 393 | CA | LEU | A | 53 | 46.646 | 22.452 | 16.826 | 1.00 | 25.70 | A | C |

FIGURE 9- 6

SUBSTITUTE SHEET (RULE 26)

16/76

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 394 | CB | LEU | A | 53 | 48.155 | 22.210 | 16.887 | 1.00 | 24.94 | A | C |
| ATOM | 395 | CG | LEU | A | 53 | 48.644 | 21.086 | 17.794 | 1.00 | 24.27 | A | C |
| ATOM | 396 | CD1 | LEU | A | 53 | 50.142 | 20.920 | 17.653 | 1.00 | 23.03 | A | C |
| ATOM | 397 | CD2 | LEU | A | 53 | 48.271 | 21.408 | 19.226 | 1.00 | 25.58 | A | C |
| ATOM | 398 | C | LEU | A | 53 | 46.380 | 23.665 | 15.953 | 1.00 | 28.37 | A | C |
| ATOM | 399 | O | LEU | A | 53 | 47.108 | 24.659 | 16.012 | 1.00 | 29.40 | A | O |
| ATOM | 400 | N | ALA | A | 54 | 45.334 | 23.580 | 15.138 | 1.00 | 30.32 | A | N |
| ATOM | 401 | CA | ALA | A | 54 | 44.975 | 24.670 | 14.237 | 1.00 | 31.21 | A | C |
| ATOM | 402 | CB | ALA | A | 54 | 43.709 | 24.301 | 13.449 | 1.00 | 31.81 | A | C |
| ATOM | 403 | C | ALA | A | 54 | 44.757 | 25.973 | 15.004 | 1.00 | 31.21 | A | C |
| ATOM | 404 | O | ALA | A | 54 | 43.897 | 26.047 | 15.881 | 1.00 | 31.23 | A | O |
| ATOM | 405 | N | GLY | A | 55 | 45.545 | 26.991 | 14.671 | 1.00 | 30.28 | A | N |
| ATOM | 406 | CA | GLY | A | 55 | 45.414 | 28.280 | 15.324 | 1.00 | 30.20 | A | C |
| ATOM | 407 | C | GLY | A | 55 | 46.478 | 28.514 | 16.367 | 1.00 | 30.23 | A | C |
| ATOM | 408 | O | GLY | A | 55 | 46.780 | 29.661 | 16.711 | 1.00 | 32.26 | A | O |
| ATOM | 409 | N | LYS | A | 56 | 47.051 | 27.427 | 16.873 | 1.00 | 29.47 | A | N |
| ATOM | 410 | CA | LYS | A | 56 | 48.091 | 27.517 | 17.896 | 1.00 | 27.92 | A | C |
| ATOM | 411 | CB | LYS | A | 56 | 48.127 | 26.219 | 18.705 | 1.00 | 27.80 | A | C |
| ATOM | 412 | CG | LYS | A | 56 | 48.816 | 26.347 | 20.042 | 1.00 | 28.80 | A | C |
| ATOM | 413 | CD | LYS | A | 56 | 47.954 | 27.118 | 21.021 | 1.00 | 26.93 | A | C |
| ATOM | 414 | CE | LYS | A | 56 | 48.641 | 27.226 | 22.373 | 1.00 | 28.62 | A | C |
| ATOM | 415 | NZ | LYS | A | 56 | 47.749 | 27.826 | 23.385 | 1.00 | 29.53 | A | N |
| ATOM | 416 | C | LYS | A | 56 | 49.464 | 27.762 | 17.258 | 1.00 | 27.62 | A | C |
| ATOM | 417 | O | LYS | A | 56 | 49.773 | 27.193 | 16.205 | 1.00 | 26.37 | A | O |
| ATOM | 418 | N | ALA | A | 57 | 50.287 | 28.592 | 17.895 | 1.00 | 25.66 | A | N |
| ATOM | 419 | CA | ALA | A | 57 | 51.610 | 28.898 | 17.364 | 1.00 | 25.87 | A | C |
| ATOM | 420 | CB | ALA | A | 57 | 52.103 | 30.254 | 17.910 | 1.00 | 26.25 | A | C |
| ATOM | 421 | C | ALA | A | 57 | 52.646 | 27.806 | 17.650 | 1.00 | 25.60 | A | C |
| ATOM | 422 | O | ALA | A | 57 | 53.644 | 28.041 | 18.325 | 1.00 | 24.35 | A | O |
| ATOM | 423 | N | ILE | A | 58 | 52.391 | 26.611 | 17.124 | 1.00 | 26.25 | A | N |
| ATOM | 424 | CA | ILE | A | 58 | 53.284 | 25.460 | 17.278 | 1.00 | 25.96 | A | C |
| ATOM | 425 | CB | ILE | A | 58 | 52.657 | 24.376 | 18.188 | 1.00 | 24.80 | A | C |
| ATOM | 426 | CG2 | ILE | A | 58 | 53.594 | 23.175 | 18.298 | 1.00 | 22.13 | A | C |
| ATOM | 427 | CG1 | ILE | A | 58 | 52.361 | 24.948 | 19.573 | 1.00 | 24.95 | A | C |
| ATOM | 428 | CD1 | ILE | A | 58 | 51.652 | 23.959 | 20.484 | 1.00 | 24.00 | A | C |
| ATOM | 429 | C | ILE | A | 58 | 53.502 | 24.833 | 15.893 | 1.00 | 26.05 | A | C |
| ATOM | 430 | O | ILE | A | 58 | 52.539 | 24.563 | 15.179 | 1.00 | 25.73 | A | O |
| ATOM | 431 | N | ASP | A | 59 | 54.752 | 24.614 | 15.497 | 1.00 | 25.98 | A | N |
| ATOM | 432 | CA | ASP | A | 59 | 55.004 | 23.990 | 14.201 | 1.00 | 26.89 | A | C |
| ATOM | 433 | CB | ASP | A | 59 | 56.322 | 24.465 | 13.606 | 1.00 | 27.67 | A | C |
| ATOM | 434 | CG | ASP | A | 59 | 56.263 | 25.893 | 13.165 | 1.00 | 29.93 | A | C |
| ATOM | 435 | OD1 | ASP | A | 59 | 55.302 | 26.250 | 12.455 | 1.00 | 34.11 | A | O |
| ATOM | 436 | OD2 | ASP | A | 59 | 57.170 | 26.656 | 13.521 | 1.00 | 33.21 | A | O |
| ATOM | 437 | C | ASP | A | 59 | 55.033 | 22.477 | 14.333 | 1.00 | 25.91 | A | C |
| ATOM | 438 | O | ASP | A | 59 | 55.475 | 21.940 | 15.343 | 1.00 | 24.36 | A | O |
| ATOM | 439 | N | VAL | A | 60 | 54.546 | 21.791 | 13.311 | 1.00 | 25.82 | A | N |
| ATOM | 440 | CA | VAL | A | 60 | 54.534 | 20.341 | 13.334 | 1.00 | 25.21 | A | C |
| ATOM | 441 | CB | VAL | A | 60 | 53.136 | 19.804 | 12.998 | 1.00 | 25.57 | A | C |
| ATOM | 442 | CG1 | VAL | A | 60 | 53.157 | 18.280 | 12.957 | 1.00 | 25.92 | A | C |
| ATOM | 443 | CG2 | VAL | A | 60 | 52.130 | 20.303 | 14.032 | 1.00 | 25.79 | A | C |
| ATOM | 444 | C | VAL | A | 60 | 55.545 | 19.804 | 12.329 | 1.00 | 25.49 | A | C |
| ATOM | 445 | O | VAL | A | 60 | 55.601 | 20.256 | 11.183 | 1.00 | 23.75 | A | O |
| ATOM | 446 | N | GLY | A | 61 | 56.355 | 18.853 | 12.778 | 1.00 | 26.73 | A | N |
| ATOM | 447 | CA | GLY | A | 61 | 57.353 | 18.249 | 11.913 | 1.00 | 26.23 | A | C |
| ATOM | 448 | C | GLY | A | 61 | 57.308 | 16.738 | 12.046 | 1.00 | 26.06 | A | C |
| ATOM | 449 | O | GLY | A | 61 | 56.562 | 16.213 | 12.866 | 1.00 | 26.38 | A | O |
| ATOM | 450 | N | ALA | A | 62 | 58.098 | 16.036 | 11.240 | 1.00 | 26.07 | A | N |
| ATOM | 451 | CA | ALA | A | 62 | 58.135 | 14.580 | 11.286 | 1.00 | 23.16 | A | C |
| ATOM | 452 | CB | ALA | A | 62 | 58.059 | 14.017 | 9.877 | 1.00 | 23.35 | A | C |
| ATOM | 453 | C | ALA | A | 62 | 59.406 | 14.093 | 11.979 | 1.00 | 23.28 | A | C |
| ATOM | 454 | O | ALA | A | 62 | 60.355 | 14.849 | 12.168 | 1.00 | 22.90 | A | O |
| ATOM | 455 | N | GLN | A | 63 | 59.413 | 12.823 | 12.362 | 1.00 | 23.25 | A | N |
| ATOM | 456 | CA | GLN | A | 63 | 60.554 | 12.227 | 13.036 | 1.00 | 22.38 | A | C |
| ATOM | 457 | CB | GLN | A | 63 | 60.067 | 11.157 | 14.013 | 1.00 | 20.35 | A | C |
| ATOM | 458 | CG | GLN | A | 63 | 59.297 | 11.681 | 15.212 | 1.00 | 20.79 | A | C |
| ATOM | 459 | CD | GLN | A | 63 | 58.715 | 10.552 | 16.046 | 1.00 | 22.65 | A | C |

FIGURE 9 - 7

SUBSTITUTE SHEET (RULE 26)

17/76

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 460 | OE1 | GLN | A | 63 | 59.292 | 9.466 | 16.123 | 1.00 | 19.63 | | |
| ATOM | 461 | NE2 | GLN | A | 63 | 57.571 | 10.807 | 16.684 | 1.00 | 21.92 | A | O |
| ATOM | 462 | C | GLN | A | 63 | 61.512 | 11.599 | 12.029 | 1.00 | 23.34 | A | N |
| ATOM | 463 | O | GLN | A | 63 | 62.659 | 11.273 | 12.359 | 1.00 | 23.09 | A | C |
| ATOM | 464 | N | ASN | A | 64 | 61.026 | 11.434 | 10.803 | 1.00 | 23.88 | A | O |
| ATOM | 465 | CA | ASN | A | 64 | 61.801 | 10.831 | 9.725 | 1.00 | 24.44 | A | N |
| ATOM | 466 | CB | ASN | A | 64 | 61.890 | 9.319 | 9.956 | 1.00 | 25.57 | A | C |
| ATOM | 467 | CG | ASN | A | 64 | 62.719 | 8.606 | 8.904 | 1.00 | 25.55 | A | C |
| ATOM | 468 | OD1 | ASN | A | 64 | 63.854 | 8.974 | 8.630 | 1.00 | 23.38 | A | C |
| ATOM | 469 | ND2 | ASN | A | 64 | 62.149 | 7.567 | 8.323 | 1.00 | 26.00 | A | O |
| ATOM | 470 | C | ASN | A | 64 | 61.158 | 11.123 | 8.368 | 1.00 | 24.39 | A | N |
| ATOM | 471 | O | ASN | A | 64 | 59.980 | 11.479 | 8.292 | 1.00 | 23.70 | A | C |
| ATOM | 472 | N | SER | A | 65 | 61.955 | 11.011 | 7.308 | 1.00 | 24.04 | A | O |
| ATOM | 473 | CA | SER | A | 65 | 61.486 | 11.224 | 5.941 | 1.00 | 22.98 | A | N |
| ATOM | 474 | CB | SER | A | 65 | 61.527 | 12.716 | 5.563 | 1.00 | 22.99 | A | C |
| ATOM | 475 | OG | SER | A | 65 | 62.804 | 13.286 | 5.775 | 1.00 | 25.38 | A | C |
| ATOM | 476 | C | SER | A | 65 | 62.329 | 10.383 | 4.978 | 1.00 | 21.52 | A | O |
| ATOM | 477 | O | SER | A | 65 | 63.399 | 9.899 | 5.342 | 1.00 | 22.53 | A | C |
| ATOM | 478 | N | ALA | A | 66 | 61.839 | 10.195 | 3.754 | 1.00 | 21.07 | A | O |
| ATOM | 479 | CA | ALA | A | 66 | 62.537 | 9.367 | 2.765 | 1.00 | 20.52 | A | N |
| ATOM | 480 | CB | ALA | A | 66 | 61.538 | 8.902 | 1.684 | 1.00 | 21.86 | A | C |
| ATOM | 481 | C | ALA | A | 66 | 63.767 | 9.994 | 2.100 | 1.00 | 20.09 | A | C |
| ATOM | 482 | O | ALA | A | 66 | 63.964 | 11.211 | 2.129 | 1.00 | 19.65 | A | O |
| ATOM | 483 | N | VAL | A | 67 | 64.591 | 9.140 | 1.499 | 1.00 | 20.22 | A | N |
| ATOM | 484 | CA | VAL | A | 67 | 65.792 | 9.581 | 0.804 | 1.00 | 21.42 | A | C |
| ATOM | 485 | CB | VAL | A | 67 | 66.744 | 8.396 | 0.531 | 1.00 | 21.63 | A | C |
| ATOM | 486 | CG1 | VAL | A | 67 | 67.335 | 7.894 | 1.832 | 1.00 | 19.64 | A | C |
| ATOM | 487 | CG2 | VAL | A | 67 | 66.001 | 7.283 | -0.175 | 1.00 | 18.78 | A | C |
| ATOM | 488 | C | VAL | A | 67 | 65.453 | 10.263 | -0.527 | 1.00 | 22.32 | A | C |
| ATOM | 489 | O | VAL | A | 67 | 66.295 | 10.946 | -1.107 | 1.00 | 23.21 | A | O |
| ATOM | 490 | N | GLU | A | 68 | 64.225 | 10.070 | -1.007 | 1.00 | 22.38 | A | N |
| ATOM | 491 | CA | GLU | A | 68 | 63.771 | 10.681 | -2.261 | 1.00 | 25.20 | A | C |
| ATOM | 492 | CB | GLU | A | 68 | 63.180 | 9.615 | -3.203 | 1.00 | 26.65 | A | C |
| ATOM | 493 | CG | GLU | A | 68 | 64.239 | 8.682 | -3.815 | 1.00 | 29.97 | A | C |
| ATOM | 494 | CD | GLU | A | 68 | 63.664 | 7.402 | -4.441 | 1.00 | 32.45 | A | C |
| ATOM | 495 | OE1 | GLU | A | 68 | 62.452 | 7.358 | -4.744 | 1.00 | 33.80 | A | O |
| ATOM | 496 | OE2 | GLU | A | 68 | 64.438 | 6.432 | -4.652 | 1.00 | 33.75 | A | O |
| ATOM | 497 | C | GLU | A | 68 | 62.734 | 11.753 | -1.932 | 1.00 | 24.72 | A | C |
| ATOM | 498 | O | GLU | A | 68 | 61.895 | 11.573 | -1.056 | 1.00 | 25.16 | A | O |
| ATOM | 499 | N | PRO | A | 69 | 62.786 | 12.893 | -2.627 | 1.00 | 24.68 | A | N |
| ATOM | 500 | CD | PRO | A | 69 | 63.840 | 13.291 | -3.577 | 1.00 | 22.08 | A | C |
| ATOM | 501 | CA | PRO | A | 69 | 61.842 | 13.994 | -2.381 | 1.00 | 24.13 | A | C |
| ATOM | 502 | CB | PRO | A | 69 | 62.509 | 15.173 | -3.081 | 1.00 | 24.61 | A | C |
| ATOM | 503 | CG | PRO | A | 69 | 63.240 | 14.507 | -4.231 | 1.00 | 21.89 | A | C |
| ATOM | 504 | C | PRO | A | 69 | 60.366 | 13.853 | -2.782 | 1.00 | 24.01 | A | C |
| ATOM | 505 | O | PRO | A | 69 | 59.476 | 14.236 | -2.022 | 1.00 | 22.76 | A | O |
| ATOM | 506 | N | MET | A | 70 | 60.104 | 13.308 | -3.964 | 1.00 | 24.45 | A | N |
| ATOM | 507 | CA | MET | A | 70 | 58.734 | 13.184 | -4.454 | 1.00 | 24.73 | A | C |
| ATOM | 508 | CB | MET | A | 70 | 58.705 | 13.542 | -5.942 | 1.00 | 25.11 | A | C |
| ATOM | 509 | CG | MET | A | 70 | 59.382 | 14.867 | -6.266 | 1.00 | 25.32 | A | C |
| ATOM | 510 | SD | MET | A | 70 | 58.694 | 16.222 | -5.341 | 1.00 | 22.46 | A | S |
| ATOM | 511 | CE | MET | A | 70 | 57.273 | 16.597 | -6.327 | 1.00 | 24.68 | A | C |
| ATOM | 512 | C | MET | A | 70 | 58.068 | 11.824 | -4.248 | 1.00 | 24.57 | A | C |
| ATOM | 513 | O | MET | A | 70 | 58.725 | 10.848 | -3.898 | 1.00 | 22.32 | A | O |
| ATOM | 514 | N | GLN | A | 71 | 56.755 | 11.778 | -4.465 | 1.00 | 23.73 | A | N |
| ATOM | 515 | CA | GLN | A | 71 | 55.998 | 10.537 | -4.331 | 1.00 | 26.34 | A | C |
| ATOM | 516 | CB | GLN | A | 71 | 54.521 | 10.766 | -4.672 | 1.00 | 26.99 | A | C |
| ATOM | 517 | CG | GLN | A | 71 | 53.848 | 11.859 | -3.841 | 1.00 | 27.16 | A | C |
| ATOM | 518 | CD | GLN | A | 71 | 52.355 | 11.970 | -4.108 | 1.00 | 28.65 | A | C |
| ATOM | 519 | OE1 | GLN | A | 71 | 51.752 | 13.031 | -3.907 | 1.00 | 31.52 | A | O |
| ATOM | 520 | NE2 | GLN | A | 71 | 51.744 | 10.871 | -4.546 | 1.00 | 27.09 | A | N |
| ATOM | 521 | C | GLN | A | 71 | 56.605 | 9.524 | -5.297 | 1.00 | 26.14 | A | C |
| ATOM | 522 | O | GLN | A | 71 | 57.031 | 9.886 | -6.390 | 1.00 | 25.94 | A | O |
| ATOM | 523 | N | GLY | A | 72 | 56.659 | 8.258 | -4.899 | 1.00 | 27.89 | A | N |
| ATOM | 524 | CA | GLY | A | 72 | 57.249 | 7.264 | -5.781 | 1.00 | 27.30 | A | C |
| ATOM | 525 | C | GLY | A | 72 | 57.155 | 5.834 | -5.294 | 1.00 | 28.77 | A | C |

FIGURE 9 - 8

SUBSTITUTE SHEET (RULE 26)

18/76

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 526 | O | GLY | A | 72 | 56.314 | 5.511 | -4.450 | 1.00 | 28.92 | A | O |
| ATOM | 527 | N | ALA | A | 73 | 58.030 | 4.982 | -5.824 | 1.00 | 28.75 | A | N |
| ATOM | 528 | CA | ALA | A | 73 | 58.042 | 3.566 | -5.482 | 1.00 | 29.81 | A | C |
| ATOM | 529 | CB | ALA | A | 73 | 58.324 | 2.742 | -6.736 | 1.00 | 29.64 | A | C |
| ATOM | 530 | C | ALA | A | 73 | 58.980 | 3.124 | -4.359 | 1.00 | 30.03 | A | C |
| ATOM | 531 | O | ALA | A | 73 | 59.935 | 2.369 | -4.589 | 1.00 | 30.59 | A | O |
| ATOM | 532 | N | LEU | A | 74 | 58.715 | 3.591 | -3.143 | 1.00 | 29.48 | A | N |
| ATOM | 533 | CA | LEU | A | 74 | 59.508 | 3.188 | -1.986 | 1.00 | 28.80 | A | C |
| ATOM | 534 | CB | LEU | A | 74 | 60.400 | 4.334 | -1.493 | 1.00 | 28.10 | A | C |
| ATOM | 535 | CG | LEU | A | 74 | 61.759 | 4.578 | -2.174 | 1.00 | 29.21 | A | C |
| ATOM | 536 | CD1 | LEU | A | 74 | 62.498 | 5.675 | -1.403 | 1.00 | 27.99 | A | C |
| ATOM | 537 | CD2 | LEU | A | 74 | 62.614 | 3.309 | -2.181 | 1.00 | 24.16 | A | C |
| ATOM | 538 | C | LEU | A | 74 | 58.527 | 2.756 | -0.892 | 1.00 | 28.20 | A | C |
| ATOM | 539 | O | LEU | A | 74 | 58.114 | 3.556 | -0.048 | 1.00 | 28.41 | A | O |
| ATOM | 540 | N | THR | A | 75 | 58.143 | 1.482 | -0.934 | 1.00 | 27.36 | A | N |
| ATOM | 541 | CA | THR | A | 75 | 57.196 | 0.924 | 0.019 | 1.00 | 25.34 | A | C |
| ATOM | 542 | CB | THR | A | 75 | 57.040 | -0.597 | -0.158 | 1.00 | 25.81 | A | C |
| ATOM | 543 | OG1 | THR | A | 75 | 56.560 | -0.884 | -1.478 | 1.00 | 26.58 | A | O |
| ATOM | 544 | CG2 | THR | A | 75 | 56.050 | -1.141 | 0.853 | 1.00 | 25.10 | A | C |
| ATOM | 545 | C | THR | A | 75 | 57.567 | 1.192 | 1.469 | 1.00 | 25.01 | A | C |
| ATOM | 546 | O | THR | A | 75 | 58.694 | 0.909 | 1.908 | 1.00 | 24.11 | A | O |
| ATOM | 547 | N | GLY | A | 76 | 56.604 | 1.742 | 2.202 | 1.00 | 22.60 | A | N |
| ATOM | 548 | CA | GLY | A | 76 | 56.807 | 2.034 | 3.606 | 1.00 | 21.52 | A | C |
| ATOM | 549 | C | GLY | A | 76 | 57.516 | 3.336 | 3.931 | 1.00 | 21.66 | A | C |
| ATOM | 550 | O | GLY | A | 76 | 57.773 | 3.609 | 5.103 | 1.00 | 21.16 | A | O |
| ATOM | 551 | N | GLU | A | 77 | 57.831 | 4.141 | 2.917 | 1.00 | 22.09 | A | N |
| ATOM | 552 | CA | GLU | A | 77 | 58.519 | 5.411 | 3.152 | 1.00 | 22.69 | A | C |
| ATOM | 553 | CB | GLU | A | 77 | 59.735 | 5.524 | 2.235 | 1.00 | 23.74 | A | C |
| ATOM | 554 | CG | GLU | A | 77 | 60.800 | 4.484 | 2.511 | 1.00 | 24.62 | A | C |
| ATOM | 555 | CD | GLU | A | 77 | 61.387 | 4.615 | 3.889 | 1.00 | 26.31 | A | C |
| ATOM | 556 | OE1 | GLU | A | 77 | 61.324 | 3.631 | 4.664 | 1.00 | 29.16 | A | O |
| ATOM | 557 | OE2 | GLU | A | 77 | 61.917 | 5.703 | 4.199 | 1.00 | 26.87 | A | O |
| ATOM | 558 | C | GLU | A | 77 | 57.622 | 6.629 | 2.959 | 1.00 | 22.38 | A | C |
| ATOM | 559 | O | GLU | A | 77 | 56.618 | 6.559 | 2.265 | 1.00 | 23.60 | A | O |
| ATOM | 560 | N | THR | A | 78 | 57.993 | 7.748 | 3.573 | 1.00 | 21.72 | A | N |
| ATOM | 561 | CA | THR | A | 78 | 57.204 | 8.964 | 3.449 | 1.00 | 21.90 | A | C |
| ATOM | 562 | CB | THR | A | 78 | 56.709 | 9.455 | 4.820 | 1.00 | 23.71 | A | C |
| ATOM | 563 | OG1 | THR | A | 78 | 55.855 | 8.465 | 5.406 | 1.00 | 26.17 | A | O |
| ATOM | 564 | CG2 | THR | A | 78 | 55.940 | 10.768 | 4.666 | 1.00 | 24.14 | A | C |
| ATOM | 565 | C | THR | A | 78 | 57.966 | 10.107 | 2.791 | 1.00 | 21.67 | A | C |
| ATOM | 566 | O | THR | A | 78 | 58.921 | 10.640 | 3.359 | 1.00 | 21.48 | A | O |
| ATOM | 567 | N | ALA | A | 79 | 57.523 | 10.503 | 1.604 | 1.00 | 21.05 | A | N |
| ATOM | 568 | CA | ALA | A | 79 | 58.186 | 11.581 | 0.887 | 1.00 | 20.84 | A | C |
| ATOM | 569 | CB | ALA | A | 79 | 57.632 | 11.692 | -0.512 | 1.00 | 21.99 | A | C |
| ATOM | 570 | C | ALA | A | 79 | 58.067 | 12.926 | 1.587 | 1.00 | 22.10 | A | C |
| ATOM | 571 | O | ALA | A | 79 | 56.991 | 13.311 | 2.046 | 1.00 | 20.64 | A | O |
| ATOM | 572 | N | PRO | A | 80 | 59.181 | 13.662 | 1.688 | 1.00 | 22.73 | A | N |
| ATOM | 573 | CD | PRO | A | 80 | 60.585 | 13.349 | 1.375 | 1.00 | 21.40 | A | C |
| ATOM | 574 | CA | PRO | A | 80 | 59.066 | 14.960 | 2.354 | 1.00 | 22.76 | A | C |
| ATOM | 575 | CB | PRO | A | 80 | 60.509 | 15.490 | 2.356 | 1.00 | 22.32 | A | C |
| ATOM | 576 | CG | PRO | A | 80 | 61.179 | 14.735 | 1.231 | 1.00 | 22.88 | A | C |
| ATOM | 577 | C | PRO | A | 80 | 58.070 | 15.881 | 1.641 | 1.00 | 23.10 | A | C |
| ATOM | 578 | O | PRO | A | 80 | 57.491 | 16.765 | 2.263 | 1.00 | 23.07 | A | O |
| ATOM | 579 | N | SER | A | 81 | 57.861 | 15.677 | 0.342 | 1.00 | 24.29 | A | N |
| ATOM | 580 | CA | SER | A | 81 | 56.893 | 16.508 | -0.372 | 1.00 | 24.61 | A | C |
| ATOM | 581 | CB | SER | A | 81 | 56.933 | 16.253 | -1.885 | 1.00 | 25.30 | A | C |
| ATOM | 582 | OG | SER | A | 81 | 56.484 | 14.951 | -2.214 | 1.00 | 29.68 | A | O |
| ATOM | 583 | C | SER | A | 81 | 55.489 | 16.210 | 0.181 | 1.00 | 23.50 | A | C |
| ATOM | 584 | O | SER | A | 81 | 54.604 | 17.065 | 0.150 | 1.00 | 23.06 | A | O |
| ATOM | 585 | N | GLN | A | 82 | 55.285 | 14.995 | 0.680 | 1.00 | 22.16 | A | N |
| ATOM | 586 | CA | GLN | A | 82 | 53.994 | 14.651 | 1.266 | 1.00 | 22.36 | A | C |
| ATOM | 587 | CB | GLN | A | 82 | 53.846 | 13.135 | 1.416 | 1.00 | 22.42 | A | C |
| ATOM | 588 | CG | GLN | A | 82 | 53.635 | 12.394 | 0.094 | 1.00 | 24.15 | A | C |
| ATOM | 589 | CD | GLN | A | 82 | 53.498 | 10.889 | 0.283 | 1.00 | 23.41 | A | C |
| ATOM | 590 | OE1 | GLN | A | 82 | 54.399 | 10.229 | 0.807 | 1.00 | 24.37 | A | O |
| ATOM | 591 | NE2 | GLN | A | 82 | 52.367 | 10.340 | -0.147 | 1.00 | 24.68 | A | N |

FIGURE 9 - 9

SUBSTITUTE SHEET (RULE 26)

19/76

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 592 | C | GLN | A | 82 | 53.866 | 15.330 | 2.635 | 1.00 | 21.01 | A | C |
| ATOM | 593 | O | GLN | A | 82 | 52.788 | 15.772 | 3.007 | 1.00 | 22.59 | A | O |
| ATOM | 594 | N | LEU | A | 83 | 54.967 | 15.411 | 3.382 | 1.00 | 20.73 | A | N |
| ATOM | 595 | CA | LEU | A | 83 | 54.948 | 16.060 | 4.696 | 1.00 | 22.14 | A | C |
| ATOM | 596 | CB | LEU | A | 83 | 56.272 | 15.832 | 5.428 | 1.00 | 21.53 | A | C |
| ATOM | 597 | CG | LEU | A | 83 | 56.523 | 14.397 | 5.923 | 1.00 | 20.86 | A | C |
| ATOM | 598 | CD1 | LEU | A | 83 | 57.922 | 14.286 | 6.498 | 1.00 | 20.79 | A | C |
| ATOM | 599 | CD2 | LEU | A | 83 | 55.489 | 14.015 | 6.965 | 1.00 | 18.53 | A | C |
| ATOM | 600 | C | LEU | A | 83 | 54.664 | 17.563 | 4.569 | 1.00 | 23.42 | A | C |
| ATOM | 601 | O | LEU | A | 83 | 53.933 | 18.128 | 5.383 | 1.00 | 24.38 | A | O |
| ATOM | 602 | N | ALA | A | 84 | 55.240 | 18.199 | 3.547 | 1.00 | 23.55 | A | N |
| ATOM | 603 | CA | ALA | A | 84 | 55.022 | 19.621 | 3.294 | 1.00 | 24.17 | A | C |
| ATOM | 604 | CB | ALA | A | 84 | 55.908 | 20.099 | 2.144 | 1.00 | 20.88 | A | C |
| ATOM | 605 | C | ALA | A | 84 | 53.548 | 19.823 | 2.937 | 1.00 | 24.79 | A | C |
| ATOM | 606 | O | ALA | A | 84 | 52.893 | 20.723 | 3.458 | 1.00 | 23.13 | A | O |
| ATOM | 607 | N | ASP | A | 85 | 53.038 | 18.965 | 2.054 | 1.00 | 26.35 | A | N |
| ATOM | 608 | CA | ASP | A | 85 | 51.644 | 19.015 | 1.620 | 1.00 | 29.13 | A | C |
| ATOM | 609 | CB | ASP | A | 85 | 51.335 | 17.834 | 0.693 | 1.00 | 32.58 | A | C |
| ATOM | 610 | CG | ASP | A | 85 | 51.834 | 18.050 | -0.727 | 1.00 | 38.34 | A | C |
| ATOM | 611 | OD1 | ASP | A | 85 | 51.920 | 17.050 | -1.488 | 1.00 | 40.17 | A | O |
| ATOM | 612 | OD2 | ASP | A | 85 | 52.126 | 19.216 | -1.093 | 1.00 | 39.51 | A | O |
| ATOM | 613 | C | ASP | A | 85 | 50.622 | 19.011 | 2.762 | 1.00 | 28.38 | A | C |
| ATOM | 614 | O | ASP | A | 85 | 49.621 | 19.714 | 2.679 | 1.00 | 27.74 | A | O |
| ATOM | 615 | N | VAL | A | 86 | 50.848 | 18.209 | 3.805 | 1.00 | 27.03 | A | N |
| ATOM | 616 | CA | VAL | A | 86 | 49.901 | 18.151 | 4.921 | 1.00 | 26.68 | A | C |
| ATOM | 617 | CB | VAL | A | 86 | 49.909 | 16.781 | 5.651 | 1.00 | 26.23 | A | C |
| ATOM | 618 | CG1 | VAL | A | 86 | 49.401 | 15.700 | 4.734 | 1.00 | 26.80 | A | C |
| ATOM | 619 | CG2 | VAL | A | 86 | 51.301 | 16.462 | 6.159 | 1.00 | 23.27 | A | C |
| ATOM | 620 | C | VAL | A | 86 | 50.107 | 19.220 | 5.985 | 1.00 | 26.52 | A | C |
| ATOM | 621 | O | VAL | A | 86 | 49.337 | 19.293 | 6.938 | 1.00 | 27.66 | A | O |
| ATOM | 622 | N | GLY | A | 87 | 51.153 | 20.026 | 5.845 | 1.00 | 26.80 | A | N |
| ATOM | 623 | CA | GLY | A | 87 | 51.391 | 21.082 | 6.812 | 1.00 | 27.37 | A | C |
| ATOM | 624 | C | GLY | A | 87 | 52.650 | 21.014 | 7.661 | 1.00 | 28.04 | A | C |
| ATOM | 625 | O | GLY | A | 87 | 52.842 | 21.867 | 8.522 | 1.00 | 29.78 | A | O |
| ATOM | 626 | N | CYS | A | 88 | 53.504 | 20.018 | 7.453 | 1.00 | 27.54 | A | N |
| ATOM | 627 | CA | CYS | A | 88 | 54.737 | 19.931 | 8.232 | 1.00 | 27.54 | A | C |
| ATOM | 628 | CB | CYS | A | 88 | 55.379 | 18.544 | 8.097 | 1.00 | 25.07 | A | C |
| ATOM | 629 | SG | CYS | A | 88 | 54.443 | 17.223 | 8.908 | 1.00 | 29.90 | A | S |
| ATOM | 630 | C | CYS | A | 88 | 55.712 | 20.991 | 7.744 | 1.00 | 27.31 | A | C |
| ATOM | 631 | O | CYS | A | 88 | 55.704 | 21.347 | 6.560 | 1.00 | 26.27 | A | O |
| ATOM | 632 | N | SER | A | 89 | 56.550 | 21.491 | 8.655 | 1.00 | 25.97 | A | N |
| ATOM | 633 | CA | SER | A | 89 | 57.535 | 22.507 | 8.304 | 1.00 | 25.57 | A | C |
| ATOM | 634 | CB | SER | A | 89 | 57.131 | 23.851 | 8.918 | 1.00 | 27.21 | A | C |
| ATOM | 635 | OG | SER | A | 89 | 56.992 | 23.755 | 10.322 | 1.00 | 29.26 | A | O |
| ATOM | 636 | C | SER | A | 89 | 58.973 | 22.152 | 8.711 | 1.00 | 24.67 | A | C |
| ATOM | 637 | O | SER | A | 89 | 59.917 | 22.824 | 8.316 | 1.00 | 23.77 | A | O |
| ATOM | 638 | N | MET | A | 90 | 59.125 | 21.092 | 9.496 | 1.00 | 25.09 | A | N |
| ATOM | 639 | CA | MET | A | 90 | 60.433 | 20.624 | 9.967 | 1.00 | 24.93 | A | C |
| ATOM | 640 | CB | MET | A | 90 | 60.676 | 21.044 | 11.424 | 1.00 | 23.18 | A | C |
| ATOM | 641 | CG | MET | A | 90 | 60.780 | 22.542 | 11.696 | 1.00 | 23.06 | A | C |
| ATOM | 642 | SD | MET | A | 90 | 60.886 | 22.917 | 13.481 | 1.00 | 17.60 | A | S |
| ATOM | 643 | CE | MET | A | 90 | 62.486 | 22.460 | 13.830 | 1.00 | 17.74 | A | C |
| ATOM | 644 | C | MET | A | 90 | 60.449 | 19.090 | 9.915 | 1.00 | 25.98 | A | C |
| ATOM | 645 | O | MET | A | 90 | 59.412 | 18.456 | 9.699 | 1.00 | 27.26 | A | O |
| ATOM | 646 | N | VAL | A | 91 | 61.619 | 18.495 | 10.118 | 1.00 | 25.13 | A | N |
| ATOM | 647 | CA | VAL | A | 91 | 61.735 | 17.041 | 10.135 | 1.00 | 23.49 | A | C |
| ATOM | 648 | CB | VAL | A | 91 | 61.593 | 16.421 | 8.715 | 1.00 | 24.95 | A | C |
| ATOM | 649 | CG1 | VAL | A | 91 | 62.724 | 16.888 | 7.814 | 1.00 | 25.63 | A | C |
| ATOM | 650 | CG2 | VAL | A | 91 | 61.591 | 14.883 | 8.813 | 1.00 | 27.62 | A | C |
| ATOM | 651 | C | VAL | A | 91 | 63.073 | 16.645 | 10.730 | 1.00 | 22.75 | A | C |
| ATOM | 652 | O | VAL | A | 91 | 64.098 | 17.256 | 10.439 | 1.00 | 22.25 | A | O |
| ATOM | 653 | N | LEU | A | 92 | 63.054 | 15.635 | 11.588 | 1.00 | 20.95 | A | N |
| ATOM | 654 | CA | LEU | A | 92 | 64.273 | 15.165 | 12.222 | 1.00 | 22.13 | A | C |
| ATOM | 655 | CB | LEU | A | 92 | 63.950 | 14.351 | 13.479 | 1.00 | 19.68 | A | C |
| ATOM | 656 | CG | LEU | A | 92 | 63.164 | 15.012 | 14.606 | 1.00 | 21.51 | A | C |
| ATOM | 657 | CD1 | LEU | A | 92 | 63.038 | 14.009 | 15.756 | 1.00 | 18.37 | A | C |

FIGURE 9- 10

SUBSTITUTE SHEET (RULE 26)

20/76

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 658 | CD2 | LEU | A | 92 | 63.867 | 16.296 | 15.068 | 1.00 | 21.83 | A | C |
| ATOM | 659 | C | LEU | A | 92 | 65.063 | 14.291 | 11.260 | 1.00 | 22.03 | A | C |
| ATOM | 660 | O | LEU | A | 92 | 64.493 | 13.583 | 10.439 | 1.00 | 21.86 | A | O |
| ATOM | 661 | N | VAL | A | 93 | 66.383 | 14.343 | 11.372 | 1.00 | 22.74 | A | N |
| ATOM | 662 | CA | VAL | A | 93 | 67.240 | 13.546 | 10.517 | 1.00 | 22.20 | A | C |
| ATOM | 663 | CB | VAL | A | 93 | 67.667 | 14.345 | 9.273 | 1.00 | 24.90 | A | C |
| ATOM | 664 | CG1 | VAL | A | 93 | 68.431 | 13.449 | 8.320 | 1.00 | 26.72 | A | C |
| ATOM | 665 | CG2 | VAL | A | 93 | 66.448 | 14.913 | 8.582 | 1.00 | 26.54 | A | C |
| ATOM | 666 | C | VAL | A | 93 | 68.484 | 13.082 | 11.272 | 1.00 | 21.59 | A | C |
| ATOM | 667 | O | VAL | A | 93 | 69.117 | 13.851 | 11.990 | 1.00 | 19.24 | A | O |
| ATOM | 668 | N | GLY | A | 94 | 68.820 | 11.808 | 11.119 | 1.00 | 21.88 | A | N |
| ATOM | 669 | CA | GLY | A | 94 | 70.000 | 11.281 | 11.780 | 1.00 | 19.89 | A | C |
| ATOM | 670 | C | GLY | A | 94 | 69.878 | 11.154 | 13.284 | 1.00 | 21.59 | A | C |
| ATOM | 671 | O | GLY | A | 94 | 70.874 | 11.263 | 13.993 | 1.00 | 20.89 | A | O |
| ATOM | 672 | N | HIS | A | 95 | 68.665 | 10.936 | 13.783 | 1.00 | 20.35 | A | N |
| ATOM | 673 | CA | HIS | A | 95 | 68.492 | 10.779 | 15.214 | 1.00 | 20.73 | A | C |
| ATOM | 674 | CB | HIS | A | 95 | 67.023 | 10.535 | 15.582 | 1.00 | 20.59 | A | C |
| ATOM | 675 | CG | HIS | A | 95 | 66.781 | 10.481 | 17.059 | 1.00 | 20.82 | A | C |
| ATOM | 676 | CD2 | HIS | A | 95 | 66.172 | 11.356 | 17.897 | 1.00 | 19.90 | A | C |
| ATOM | 677 | ND1 | HIS | A | 95 | 67.273 | 9.469 | 17.854 | 1.00 | 18.89 | A | N |
| ATOM | 678 | CE1 | HIS | A | 95 | 66.983 | 9.725 | 19.118 | 1.00 | 20.99 | A | C |
| ATOM | 679 | NE2 | HIS | A | 95 | 66.317 | 10.864 | 19.173 | 1.00 | 19.67 | A | N |
| ATOM | 680 | C | HIS | A | 95 | 69.336 | 9.583 | 15.648 | 1.00 | 20.75 | A | C |
| ATOM | 681 | O | HIS | A | 95 | 69.428 | 8.582 | 14.942 | 1.00 | 20.56 | A | O |
| ATOM | 682 | N | SER | A | 96 | 69.951 | 9.710 | 16.812 | 1.00 | 20.97 | A | N |
| ATOM | 683 | CA | SER | A | 96 | 70.808 | 8.674 | 17.369 | 1.00 | 22.86 | A | C |
| ATOM | 684 | CB | SER | A | 96 | 71.102 | 9.000 | 18.824 | 1.00 | 22.43 | A | C |
| ATOM | 685 | OG | SER | A | 96 | 71.866 | 7.970 | 19.404 | 1.00 | 32.09 | A | O |
| ATOM | 686 | C | SER | A | 96 | 70.234 | 7.260 | 17.283 | 1.00 | 23.31 | A | C |
| ATOM | 687 | O | SER | A | 96 | 70.959 | 6.310 | 17.012 | 1.00 | 23.47 | A | O |
| ATOM | 688 | N | GLU | A | 97 | 68.933 | 7.131 | 17.526 | 1.00 | 23.37 | A | N |
| ATOM | 689 | CA | GLU | A | 97 | 68.261 | 5.842 | 17.480 | 1.00 | 24.40 | A | C |
| ATOM | 690 | CB | GLU | A | 97 | 66.821 | 5.991 | 18.008 | 1.00 | 24.48 | A | C |
| ATOM | 691 | CG | GLU | A | 97 | 66.736 | 6.016 | 19.544 | 1.00 | 25.49 | A | C |
| ATOM | 692 | CD | GLU | A | 97 | 65.596 | 6.881 | 20.109 | 1.00 | 27.14 | A | C |
| ATOM | 693 | OE1 | GLU | A | 97 | 64.545 | 7.030 | 19.441 | 1.00 | 26.09 | A | O |
| ATOM | 694 | OE2 | GLU | A | 97 | 65.755 | 7.396 | 21.245 | 1.00 | 27.01 | A | O |
| ATOM | 695 | C | GLU | A | 97 | 68.272 | 5.237 | 16.075 | 1.00 | 24.38 | A | C |
| ATOM | 696 | O | GLU | A | 97 | 68.481 | 4.036 | 15.913 | 1.00 | 25.68 | A | O |
| ATOM | 697 | N | ARG | A | 98 | 68.066 | 6.059 | 15.055 | 1.00 | 23.13 | A | N |
| ATOM | 698 | CA | ARG | A | 98 | 68.078 | 5.536 | 13.694 | 1.00 | 23.18 | A | C |
| ATOM | 699 | CB | ARG | A | 98 | 67.498 | 6.566 | 12.727 | 1.00 | 20.73 | A | C |
| ATOM | 700 | CG | ARG | A | 98 | 66.060 | 6.970 | 13.051 | 1.00 | 18.77 | A | C |
| ATOM | 701 | CD | ARG | A | 98 | 65.394 | 7.645 | 11.857 | 1.00 | 18.93 | A | C |
| ATOM | 702 | NE | ARG | A | 98 | 65.158 | 6.711 | 10.759 | 1.00 | 17.29 | A | N |
| ATOM | 703 | CZ | ARG | A | 98 | 64.170 | 5.818 | 10.725 | 1.00 | 18.68 | A | C |
| ATOM | 704 | NH1 | ARG | A | 98 | 63.306 | 5.727 | 11.729 | 1.00 | 17.80 | A | N |
| ATOM | 705 | NH2 | ARG | A | 98 | 64.048 | 5.005 | 9.680 | 1.00 | 17.77 | A | N |
| ATOM | 706 | C | ARG | A | 98 | 69.495 | 5.128 | 13.262 | 1.00 | 25.05 | A | C |
| ATOM | 707 | O | ARG | A | 98 | 69.680 | 4.106 | 12.585 | 1.00 | 25.50 | A | O |
| ATOM | 708 | N | ARG | A | 99 | 70.491 | 5.915 | 13.661 | 1.00 | 24.25 | A | N |
| ATOM | 709 | CA | ARG | A | 99 | 71.885 | 5.628 | 13.316 | 1.00 | 25.21 | A | C |
| ATOM | 710 | CB | ARG | A | 99 | 72.797 | 6.789 | 13.739 | 1.00 | 25.41 | A | C |
| ATOM | 711 | CG | ARG | A | 99 | 72.724 | 8.067 | 12.910 | 1.00 | 24.27 | A | C |
| ATOM | 712 | CD | ARG | A | 99 | 73.656 | 9.106 | 13.529 | 1.00 | 22.41 | A | C |
| ATOM | 713 | NE | ARG | A | 99 | 73.516 | 10.446 | 12.966 | 1.00 | 20.97 | A | N |
| ATOM | 714 | CZ | ARG | A | 99 | 74.212 | 10.915 | 11.935 | 1.00 | 21.14 | A | C |
| ATOM | 715 | NH1 | ARG | A | 99 | 75.109 | 10.154 | 11.331 | 1.00 | 18.62 | A | N |
| ATOM | 716 | NH2 | ARG | A | 99 | 74.021 | 12.162 | 11.523 | 1.00 | 19.85 | A | N |
| ATOM | 717 | C | ARG | A | 99 | 72.421 | 4.357 | 13.982 | 1.00 | 25.58 | A | C |
| ATOM | 718 | O | ARG | A | 99 | 72.998 | 3.490 | 13.327 | 1.00 | 26.49 | A | O |
| ATOM | 719 | N | LEU | A | 100 | 72.232 | 4.267 | 15.295 | 1.00 | 28.13 | A | N |
| ATOM | 720 | CA | LEU | A | 100 | 72.730 | 3.144 | 16.089 | 1.00 | 28.79 | A | C |
| ATOM | 721 | CB | LEU | A | 100 | 73.030 | 3.625 | 17.513 | 1.00 | 27.60 | A | C |
| ATOM | 722 | CG | LEU | A | 100 | 73.894 | 4.890 | 17.639 | 1.00 | 29.65 | A | C |
| ATOM | 723 | CD1 | LEU | A | 100 | 73.920 | 5.331 | 19.086 | 1.00 | 31.17 | A | C |

FIGURE 9 - 11

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 724 | CD2 | LEU | A | 100 | 75.313 | 4.643 | 17.135 | 1.00 | 30.34 | A | C |
| ATOM | 725 | C | LEU | A | 100 | 71.843 | 1.896 | 16.142 | 1.00 | 29.05 | A | C |
| ATOM | 726 | O | LEU | A | 100 | 72.358 | 0.778 | 16.183 | 1.00 | 29.64 | A | O |
| ATOM | 727 | N | ILE | A | 101 | 70.523 | 2.063 | 16.147 | 1.00 | 28.75 | A | N |
| ATOM | 728 | CA | ILE | A | 101 | 69.640 | 0.896 | 16.194 | 1.00 | 28.97 | A | C |
| ATOM | 729 | CB | ILE | A | 101 | 68.361 | 1.169 | 17.014 | 1.00 | 28.00 | A | C |
| ATOM | 730 | CG2 | ILE | A | 101 | 67.420 | -0.018 | 16.917 | 1.00 | 27.29 | A | C |
| ATOM | 731 | CG1 | ILE | A | 101 | 68.718 | 1.413 | 18.476 | 1.00 | 28.26 | A | C |
| ATOM | 732 | CD1 | ILE | A | 101 | 67.511 | 1.730 | 19.350 | 1.00 | 29.81 | A | C |
| ATOM | 733 | C | ILE | A | 101 | 69.215 | 0.399 | 14.814 | 1.00 | 29.21 | A | C |
| ATOM | 734 | O | ILE | A | 101 | 69.167 | -0.807 | 14.562 | 1.00 | 30.84 | A | O |
| ATOM | 735 | N | LEU | A | 102 | 68.902 | 1.324 | 13.921 | 1.00 | 29.04 | A | N |
| ATOM | 736 | CA | LEU | A | 102 | 68.465 | 0.947 | 12.588 | 1.00 | 28.21 | A | C |
| ATOM | 737 | CB | LEU | A | 102 | 67.339 | 1.882 | 12.143 | 1.00 | 26.84 | A | C |
| ATOM | 738 | CG | LEU | A | 102 | 66.193 | 1.902 | 13.163 | 1.00 | 30.37 | A | C |
| ATOM | 739 | CD1 | LEU | A | 102 | 65.140 | 2.953 | 12.797 | 1.00 | 29.87 | A | C |
| ATOM | 740 | CD2 | LEU | A | 102 | 65.574 | 0.518 | 13.230 | 1.00 | 29.13 | A | C |
| ATOM | 741 | C | LEU | A | 102 | 69.613 | 0.959 | 11.586 | 1.00 | 27.84 | A | C |
| ATOM | 742 | O | LEU | A | 102 | 69.412 | 0.697 | 10.411 | 1.00 | 27.76 | A | O |
| ATOM | 743 | N | GLY | A | 103 | 70.816 | 1.267 | 12.062 | 1.00 | 29.82 | A | N |
| ATOM | 744 | CA | GLY | A | 103 | 71.985 | 1.290 | 11.194 | 1.00 | 30.21 | A | C |
| ATOM | 745 | C | GLY | A | 103 | 71.931 | 2.246 | 10.011 | 1.00 | 30.66 | A | C |
| ATOM | 746 | O | GLY | A | 103 | 72.359 | 1.911 | 8.905 | 1.00 | 31.98 | A | O |
| ATOM | 747 | N | GLU | A | 104 | 71.422 | 3.448 | 10.243 | 1.00 | 30.08 | A | N |
| ATOM | 748 | CA | GLU | A | 104 | 71.305 | 4.460 | 9.199 | 1.00 | 28.56 | A | C |
| ATOM | 749 | CB | GLU | A | 104 | 70.196 | 5.432 | 9.600 | 1.00 | 31.15 | A | C |
| ATOM | 750 | CG | GLU | A | 104 | 69.598 | 6.253 | 8.488 | 1.00 | 32.15 | A | C |
| ATOM | 751 | CD | GLU | A | 104 | 68.150 | 6.622 | 8.785 | 1.00 | 31.78 | A | C |
| ATOM | 752 | OE1 | GLU | A | 104 | 67.270 | 5.741 | 8.687 | 1.00 | 32.80 | A | O |
| ATOM | 753 | OE2 | GLU | A | 104 | 67.894 | 7.787 | 9.132 | 1.00 | 32.07 | A | O |
| ATOM | 754 | C | GLU | A | 104 | 72.656 | 5.164 | 9.110 | 1.00 | 28.08 | A | C |
| ATOM | 755 | O | GLU | A | 104 | 73.150 | 5.679 | 10.106 | 1.00 | 26.92 | A | O |
| ATOM | 756 | N | SER | A | 105 | 73.250 | 5.192 | 7.922 | 1.00 | 27.99 | A | N |
| ATOM | 757 | CA | SER | A | 105 | 74.566 | 5.812 | 7.748 | 1.00 | 29.21 | A | C |
| ATOM | 758 | CB | SER | A | 105 | 75.332 | 5.100 | 6.631 | 1.00 | 30.66 | A | C |
| ATOM | 759 | OG | SER | A | 105 | 74.760 | 5.398 | 5.369 | 1.00 | 30.12 | A | O |
| ATOM | 760 | C | SER | A | 105 | 74.549 | 7.311 | 7.440 | 1.00 | 29.08 | A | C |
| ATOM | 761 | O | SER | A | 105 | 73.495 | 7.901 | 7.214 | 1.00 | 28.26 | A | O |
| ATOM | 762 | N | ASP | A | 106 | 75.740 | 7.909 | 7.435 | 1.00 | 29.22 | A | N |
| ATOM | 763 | CA | ASP | A | 106 | 75.913 | 9.328 | 7.145 | 1.00 | 30.37 | A | C |
| ATOM | 764 | CB | ASP | A | 106 | 77.400 | 9.709 | 7.244 | 1.00 | 32.57 | A | C |
| ATOM | 765 | CG | ASP | A | 106 | 77.952 | 9.553 | 8.674 | 1.00 | 35.47 | A | C |
| ATOM | 766 | OD1 | ASP | A | 106 | 77.221 | 9.892 | 9.636 | 1.00 | 35.84 | A | O |
| ATOM | 767 | OD2 | ASP | A | 106 | 79.110 | 9.105 | 8.842 | 1.00 | 35.34 | A | O |
| ATOM | 768 | C | ASP | A | 106 | 75.369 | 9.662 | 5.755 | 1.00 | 30.39 | A | C |
| ATOM | 769 | O | ASP | A | 106 | 74.758 | 10.716 | 5.548 | 1.00 | 30.92 | A | O |
| ATOM | 770 | N | GLU | A | 107 | 75.588 | 8.752 | 4.813 | 1.00 | 28.97 | A | N |
| ATOM | 771 | CA | GLU | A | 107 | 75.130 | 8.905 | 3.437 | 1.00 | 28.92 | A | C |
| ATOM | 772 | CB | GLU | A | 107 | 75.565 | 7.679 | 2.628 | 1.00 | 32.22 | A | C |
| ATOM | 773 | CG | GLU | A | 107 | 74.728 | 7.390 | 1.386 | 1.00 | 38.12 | A | C |
| ATOM | 774 | CD | GLU | A | 107 | 75.148 | 8.192 | 0.178 | 1.00 | 41.71 | A | C |
| ATOM | 775 | OE1 | GLU | A | 107 | 75.342 | 9.433 | 0.300 | 1.00 | 44.63 | A | O |
| ATOM | 776 | OE2 | GLU | A | 107 | 75.274 | 7.578 | -0.905 | 1.00 | 43.28 | A | O |
| ATOM | 777 | C | GLU | A | 107 | 73.604 | 9.058 | 3.402 | 1.00 | 26.24 | A | C |
| ATOM | 778 | O | GLU | A | 107 | 73.084 | 10.013 | 2.834 | 1.00 | 25.49 | A | O |
| ATOM | 779 | N | VAL | A | 108 | 72.905 | 8.105 | 4.014 | 1.00 | 24.90 | A | N |
| ATOM | 780 | CA | VAL | A | 108 | 71.443 | 8.103 | 4.093 | 1.00 | 23.42 | A | C |
| ATOM | 781 | CB | VAL | A | 108 | 70.955 | 6.902 | 4.927 | 1.00 | 22.92 | A | C |
| ATOM | 782 | CG1 | VAL | A | 108 | 69.437 | 6.964 | 5.105 | 1.00 | 23.73 | A | C |
| ATOM | 783 | CG2 | VAL | A | 108 | 71.374 | 5.604 | 4.250 | 1.00 | 25.45 | A | C |
| ATOM | 784 | C | VAL | A | 108 | 70.926 | 9.395 | 4.737 | 1.00 | 23.04 | A | C |
| ATOM | 785 | O | VAL | A | 108 | 70.012 | 10.039 | 4.219 | 1.00 | 23.34 | A | O |
| ATOM | 786 | N | VAL | A | 109 | 71.516 | 9.752 | 5.878 | 1.00 | 22.72 | A | N |
| ATOM | 787 | CA | VAL | A | 109 | 71.164 | 10.965 | 6.613 | 1.00 | 21.24 | A | C |
| ATOM | 788 | CB | VAL | A | 109 | 72.064 | 11.138 | 7.855 | 1.00 | 20.95 | A | C |
| ATOM | 789 | CG1 | VAL | A | 109 | 71.849 | 12.503 | 8.481 | 1.00 | 20.30 | A | C |

FIGURE 9- 12

| | | | | | | | | | |
|------|-----|---------------|--------|--------|--------|------|-------|---|---|
| ATOM | 790 | CG2 VAL A 109 | 71.770 | 10.041 | 8.859 | 1.00 | 19.24 | A | C |
| ATOM | 791 | C VAL A 109 | 71.354 | 12.171 | 5.707 | 1.00 | 23.17 | A | C |
| ATOM | 792 | O VAL A 109 | 70.497 | 13.058 | 5.639 | 1.00 | 22.43 | A | O |
| ATOM | 793 | N SER A 110 | 72.490 | 12.195 | 5.018 | 1.00 | 23.24 | A | N |
| ATOM | 794 | CA SER A 110 | 72.804 | 13.276 | 4.094 | 1.00 | 25.90 | A | C |
| ATOM | 795 | CB SER A 110 | 74.147 | 13.018 | 3.422 | 1.00 | 27.68 | A | C |
| ATOM | 796 | OG SER A 110 | 74.096 | 13.380 | 2.049 | 1.00 | 33.96 | A | O |
| ATOM | 797 | C SER A 110 | 71.728 | 13.373 | 3.024 | 1.00 | 25.69 | A | C |
| ATOM | 798 | O SER A 110 | 71.254 | 14.455 | 2.705 | 1.00 | 26.99 | A | O |
| ATOM | 799 | N ARG A 111 | 71.357 | 12.227 | 2.466 | 1.00 | 26.42 | A | N |
| ATOM | 800 | CA ARG A 111 | 70.337 | 12.163 | 1.427 | 1.00 | 27.72 | A | C |
| ATOM | 801 | CB ARG A 111 | 70.155 | 10.716 | 0.961 | 1.00 | 31.68 | A | C |
| ATOM | 802 | CG ARG A 111 | 71.310 | 10.144 | 0.151 | 1.00 | 35.70 | A | C |
| ATOM | 803 | CD ARG A 111 | 71.239 | 10.570 | -1.294 | 1.00 | 38.26 | A | C |
| ATOM | 804 | NE ARG A 111 | 70.045 | 10.049 | -1.964 | 1.00 | 43.07 | A | N |
| ATOM | 805 | CZ ARG A 111 | 69.804 | 8.756 | -2.184 | 1.00 | 43.65 | A | C |
| ATOM | 806 | NH1 ARG A 111 | 70.680 | 7.836 | -1.786 | 1.00 | 43.39 | A | N |
| ATOM | 807 | NH2 ARG A 111 | 68.687 | 8.385 | -2.803 | 1.00 | 42.86 | A | N |
| ATOM | 808 | C ARG A 111 | 68.990 | 12.699 | 1.915 | 1.00 | 26.74 | A | C |
| ATOM | 809 | O ARG A 111 | 68.348 | 13.493 | 1.232 | 1.00 | 26.52 | A | O |
| ATOM | 810 | N LYS A 112 | 68.565 | 12.257 | 3.095 | 1.00 | 25.15 | A | N |
| ATOM | 811 | CA LYS A 112 | 67.285 | 12.684 | 3.658 | 1.00 | 23.99 | A | C |
| ATOM | 812 | CB LYS A 112 | 66.985 | 11.875 | 4.922 | 1.00 | 22.91 | A | C |
| ATOM | 813 | CG LYS A 112 | 66.614 | 10.425 | 4.618 | 1.00 | 22.66 | A | C |
| ATOM | 814 | CD LYS A 112 | 66.452 | 9.598 | 5.873 | 1.00 | 21.50 | A | C |
| ATOM | 815 | CE LYS A 112 | 65.834 | 8.243 | 5.557 | 1.00 | 20.70 | A | C |
| ATOM | 816 | NZ LYS A 112 | 65.572 | 7.474 | 6.798 | 1.00 | 22.87 | A | N |
| ATOM | 817 | C LYS A 112 | 67.273 | 14.179 | 3.961 | 1.00 | 23.35 | A | C |
| ATOM | 818 | O LYS A 112 | 66.268 | 14.864 | 3.749 | 1.00 | 22.46 | A | O |
| ATOM | 819 | N PHE A 113 | 68.399 | 14.678 | 4.452 | 1.00 | 22.04 | A | N |
| ATOM | 820 | CA PHE A 113 | 68.522 | 16.085 | 4.764 | 1.00 | 20.99 | A | C |
| ATOM | 821 | CB PHE A 113 | 69.919 | 16.385 | 5.303 | 1.00 | 20.11 | A | C |
| ATOM | 822 | CG PHE A 113 | 70.123 | 17.818 | 5.699 | 1.00 | 20.20 | A | C |
| ATOM | 823 | CD1 PHE A 113 | 70.194 | 18.819 | 4.735 | 1.00 | 19.53 | A | C |
| ATOM | 824 | CD2 PHE A 113 | 70.253 | 18.166 | 7.036 | 1.00 | 20.00 | A | C |
| ATOM | 825 | CE1 PHE A 113 | 70.396 | 20.149 | 5.097 | 1.00 | 21.21 | A | C |
| ATOM | 826 | CE2 PHE A 113 | 70.454 | 19.495 | 7.409 | 1.00 | 22.80 | A | C |
| ATOM | 827 | CZ PHE A 113 | 70.527 | 20.489 | 6.429 | 1.00 | 20.17 | A | C |
| ATOM | 828 | C PHE A 113 | 68.264 | 16.899 | 3.507 | 1.00 | 21.21 | A | C |
| ATOM | 829 | O PHE A 113 | 67.409 | 17.788 | 3.500 | 1.00 | 20.47 | A | O |
| ATOM | 830 | N ALA A 114 | 69.005 | 16.591 | 2.445 | 1.00 | 21.11 | A | N |
| ATOM | 831 | CA ALA A 114 | 68.863 | 17.305 | 1.182 | 1.00 | 21.87 | A | C |
| ATOM | 832 | CB ALA A 114 | 69.920 | 16.825 | 0.180 | 1.00 | 22.02 | A | C |
| ATOM | 833 | C ALA A 114 | 67.464 | 17.142 | 0.596 | 1.00 | 23.19 | A | C |
| ATOM | 834 | O ALA A 114 | 66.884 | 18.097 | 0.085 | 1.00 | 23.74 | A | O |
| ATOM | 835 | N ALA A 115 | 66.916 | 15.933 | 0.667 | 1.00 | 24.08 | A | N |
| ATOM | 836 | CA ALA A 115 | 65.568 | 15.696 | 0.142 | 1.00 | 23.92 | A | C |
| ATOM | 837 | CB ALA A 115 | 65.182 | 14.231 | 0.306 | 1.00 | 23.96 | A | C |
| ATOM | 838 | C ALA A 115 | 64.544 | 16.579 | 0.857 | 1.00 | 24.23 | A | C |
| ATOM | 839 | O ALA A 115 | 63.700 | 17.219 | 0.219 | 1.00 | 24.01 | A | O |
| ATOM | 840 | N ALA A 116 | 64.623 | 16.610 | 2.183 | 1.00 | 23.50 | A | N |
| ATOM | 841 | CA ALA A 116 | 63.694 | 17.393 | 2.977 | 1.00 | 24.08 | A | C |
| ATOM | 842 | CB ALA A 116 | 63.941 | 17.144 | 4.468 | 1.00 | 22.52 | A | C |
| ATOM | 843 | C ALA A 116 | 63.817 | 18.883 | 2.652 | 1.00 | 24.86 | A | C |
| ATOM | 844 | O ALA A 116 | 62.811 | 19.574 | 2.524 | 1.00 | 23.78 | A | O |
| ATOM | 845 | N GLN A 117 | 65.047 | 19.367 | 2.504 | 1.00 | 25.50 | A | N |
| ATOM | 846 | CA GLN A 117 | 65.271 | 20.779 | 2.193 | 1.00 | 27.82 | A | C |
| ATOM | 847 | CB GLN A 117 | 66.765 | 21.152 | 2.236 | 1.00 | 27.90 | A | C |
| ATOM | 848 | CG GLN A 117 | 67.350 | 21.393 | 3.613 | 1.00 | 28.25 | A | C |
| ATOM | 849 | CD GLN A 117 | 68.505 | 22.403 | 3.595 | 1.00 | 28.24 | A | C |
| ATOM | 850 | OE1 GLN A 117 | 69.410 | 22.317 | 2.769 | 1.00 | 26.74 | A | O |
| ATOM | 851 | NE2 GLN A 117 | 68.471 | 23.357 | 4.518 | 1.00 | 25.64 | A | N |
| ATOM | 852 | C GLN A 117 | 64.738 | 21.171 | 0.830 | 1.00 | 28.02 | A | C |
| ATOM | 853 | O GLN A 117 | 64.297 | 22.298 | 0.643 | 1.00 | 27.95 | A | O |
| ATOM | 854 | N SER A 118 | 64.785 | 20.263 | -0.135 | 1.00 | 27.71 | A | N |
| ATOM | 855 | CA SER A 118 | 64.300 | 20.637 | -1.456 | 1.00 | 29.67 | A | C |

FIGURE 9- 13

23/76

| | | | | | | | | | | | | |
|------|-----|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 856 | CB | SER | A | 118 | 64.761 | 19.633 | -2.523 | 1.00 | 29.29 | A | C |
| ATOM | 857 | OG | SER | A | 118 | 64.565 | 18.302 | -2.108 | 1.00 | 32.45 | A | O |
| ATOM | 858 | C | SER | A | 118 | 62.792 | 20.799 | -1.468 | 1.00 | 29.56 | A | C |
| ATOM | 859 | O | SER | A | 118 | 62.241 | 21.429 | -2.370 | 1.00 | 31.23 | A | O |
| ATOM | 860 | N | CYS | A | 119 | 62.122 | 20.245 | -0.462 | 1.00 | 28.63 | A | N |
| ATOM | 861 | CA | CYS | A | 119 | 60.670 | 20.367 | -0.373 | 1.00 | 27.50 | A | C |
| ATOM | 862 | CB | CYS | A | 119 | 60.049 | 19.044 | 0.086 | 1.00 | 28.03 | A | C |
| ATOM | 863 | SG | CYS | A | 119 | 60.134 | 17.706 | -1.153 | 1.00 | 28.40 | A | S |
| ATOM | 864 | C | CYS | A | 119 | 60.248 | 21.508 | 0.564 | 1.00 | 27.34 | A | C |
| ATOM | 865 | O | CYS | A | 119 | 59.090 | 21.594 | 0.972 | 1.00 | 27.81 | A | O |
| ATOM | 866 | N | GLY | A | 120 | 61.192 | 22.381 | 0.894 | 1.00 | 26.88 | A | N |
| ATOM | 867 | CA | GLY | A | 120 | 60.890 | 23.508 | 1.760 | 1.00 | 27.00 | A | C |
| ATOM | 868 | C | GLY | A | 120 | 60.925 | 23.237 | 3.255 | 1.00 | 27.79 | A | C |
| ATOM | 869 | O | GLY | A | 120 | 60.708 | 24.150 | 4.051 | 1.00 | 28.94 | A | O |
| ATOM | 870 | N | LEU | A | 121 | 61.204 | 21.999 | 3.651 | 1.00 | 25.39 | A | N |
| ATOM | 871 | CA | LEU | A | 121 | 61.246 | 21.662 | 5.069 | 1.00 | 24.49 | A | C |
| ATOM | 872 | CB | LEU | A | 121 | 61.019 | 20.155 | 5.256 | 1.00 | 23.15 | A | C |
| ATOM | 873 | CG | LEU | A | 121 | 59.704 | 19.556 | 4.742 | 1.00 | 23.30 | A | C |
| ATOM | 874 | CD1 | LEU | A | 121 | 59.696 | 18.039 | 4.954 | 1.00 | 20.41 | A | C |
| ATOM | 875 | CD2 | LEU | A | 121 | 58.530 | 20.209 | 5.469 | 1.00 | 19.42 | A | C |
| ATOM | 876 | C | LEU | A | 121 | 62.572 | 22.065 | 5.726 | 1.00 | 24.61 | A | C |
| ATOM | 877 | O | LEU | A | 121 | 63.622 | 22.090 | 5.077 | 1.00 | 24.07 | A | O |
| ATOM | 878 | N | VAL | A | 122 | 62.514 | 22.390 | 7.012 | 1.00 | 23.19 | A | N |
| ATOM | 879 | CA | VAL | A | 122 | 63.710 | 22.766 | 7.760 | 1.00 | 23.80 | A | C |
| ATOM | 880 | CB | VAL | A | 122 | 63.432 | 23.914 | 8.759 | 1.00 | 24.98 | A | C |
| ATOM | 881 | CG1 | VAL | A | 122 | 64.705 | 24.256 | 9.518 | 1.00 | 24.55 | A | C |
| ATOM | 882 | CG2 | VAL | A | 122 | 62.898 | 25.132 | 8.025 | 1.00 | 22.48 | A | C |
| ATOM | 883 | C | VAL | A | 122 | 64.147 | 21.540 | 8.552 | 1.00 | 23.29 | A | C |
| ATOM | 884 | O | VAL | A | 122 | 63.507 | 21.166 | 9.529 | 1.00 | 21.70 | A | O |
| ATOM | 885 | N | PRO | A | 123 | 65.230 | 20.882 | 8.123 | 1.00 | 21.82 | A | N |
| ATOM | 886 | CD | PRO | A | 123 | 65.947 | 21.022 | 6.845 | 1.00 | 21.17 | A | C |
| ATOM | 887 | CA | PRO | A | 123 | 65.679 | 19.698 | 8.857 | 1.00 | 20.61 | A | C |
| ATOM | 888 | CB | PRO | A | 123 | 66.605 | 18.998 | 7.862 | 1.00 | 21.35 | A | C |
| ATOM | 889 | CG | PRO | A | 123 | 67.149 | 20.122 | 7.049 | 1.00 | 23.51 | A | C |
| ATOM | 890 | C | PRO | A | 123 | 66.376 | 20.001 | 10.174 | 1.00 | 20.10 | A | C |
| ATOM | 891 | O | PRO | A | 123 | 66.963 | 21.064 | 10.355 | 1.00 | 20.10 | A | O |
| ATOM | 892 | N | VAL | A | 124 | 66.273 | 19.056 | 11.098 | 1.00 | 20.26 | A | N |
| ATOM | 893 | CA | VAL | A | 124 | 66.907 | 19.151 | 12.406 | 1.00 | 20.51 | A | C |
| ATOM | 894 | CB | VAL | A | 124 | 65.894 | 18.893 | 13.565 | 1.00 | 19.48 | A | C |
| ATOM | 895 | CG1 | VAL | A | 124 | 66.619 | 18.883 | 14.900 | 1.00 | 19.14 | A | C |
| ATOM | 896 | CG2 | VAL | A | 124 | 64.805 | 19.958 | 13.571 | 1.00 | 21.02 | A | C |
| ATOM | 897 | C | VAL | A | 124 | 67.947 | 18.026 | 12.386 | 1.00 | 20.14 | A | C |
| ATOM | 898 | O | VAL | A | 124 | 67.628 | 16.872 | 12.650 | 1.00 | 20.29 | A | O |
| ATOM | 899 | N | LEU | A | 125 | 69.179 | 18.360 | 12.033 | 1.00 | 20.32 | A | N |
| ATOM | 900 | CA | LEU | A | 125 | 70.239 | 17.369 | 11.967 | 1.00 | 19.33 | A | C |
| ATOM | 901 | CB | LEU | A | 125 | 71.393 | 17.891 | 11.107 | 1.00 | 19.81 | A | C |
| ATOM | 902 | CG | LEU | A | 125 | 72.699 | 17.083 | 11.047 | 1.00 | 21.27 | A | C |
| ATOM | 903 | CD1 | LEU | A | 125 | 72.470 | 15.698 | 10.444 | 1.00 | 22.91 | A | C |
| ATOM | 904 | CD2 | LEU | A | 125 | 73.689 | 17.841 | 10.213 | 1.00 | 21.28 | A | C |
| ATOM | 905 | C | LEU | A | 125 | 70.735 | 17.042 | 13.365 | 1.00 | 19.71 | A | C |
| ATOM | 906 | O | LEU | A | 125 | 71.271 | 17.906 | 14.057 | 1.00 | 20.26 | A | O |
| ATOM | 907 | N | CYS | A | 126 | 70.535 | 15.798 | 13.787 | 1.00 | 18.36 | A | N |
| ATOM | 908 | CA | CYS | A | 126 | 70.986 | 15.359 | 15.102 | 1.00 | 19.87 | A | C |
| ATOM | 909 | CB | CYS | A | 126 | 70.046 | 14.292 | 15.665 | 1.00 | 19.95 | A | C |
| ATOM | 910 | SG | CYS | A | 126 | 68.309 | 14.799 | 15.734 | 1.00 | 25.01 | A | S |
| ATOM | 911 | C | CYS | A | 126 | 72.382 | 14.773 | 14.951 | 1.00 | 20.26 | A | C |
| ATOM | 912 | O | CYS | A | 126 | 72.636 | 14.003 | 14.019 | 1.00 | 19.41 | A | O |
| ATOM | 913 | N | VAL | A | 127 | 73.285 | 15.145 | 15.855 | 1.00 | 20.68 | A | N |
| ATOM | 914 | CA | VAL | A | 127 | 74.653 | 14.642 | 15.817 | 1.00 | 20.01 | A | C |
| ATOM | 915 | CB | VAL | A | 127 | 75.620 | 15.681 | 15.190 | 1.00 | 20.75 | A | C |
| ATOM | 916 | CG1 | VAL | A | 127 | 75.139 | 16.075 | 13.796 | 1.00 | 19.95 | A | C |
| ATOM | 917 | CG2 | VAL | A | 127 | 75.716 | 16.907 | 16.069 | 1.00 | 20.08 | A | C |
| ATOM | 918 | C | VAL | A | 127 | 75.087 | 14.333 | 17.241 | 1.00 | 21.23 | A | C |
| ATOM | 919 | O | VAL | A | 127 | 74.566 | 14.922 | 18.188 | 1.00 | 22.46 | A | O |
| ATOM | 920 | N | GLY | A | 128 | 76.026 | 13.400 | 17.397 | 1.00 | 19.45 | A | N |
| ATOM | 921 | CA | GLY | A | 128 | 76.496 | 13.042 | 18.726 | 1.00 | 20.51 | A | C |

FIGURE 9 - 14

SUBSTITUTE SHEET (RULE 26)

24/76

| | | | | | | | | | | |
|------|-----|-----|-----------|--------|--------|--------|------|-------|---|---|
| ATOM | 922 | C | GLY A 128 | 77.455 | 11.862 | 18.735 | 1.00 | 21.45 | A | C |
| ATOM | 923 | O | GLY A 128 | 77.491 | 11.093 | 17.777 | 1.00 | 20.63 | A | O |
| ATOM | 924 | N | GLU A 129 | 78.221 | 11.712 | 19.815 | 1.00 | 21.31 | A | N |
| ATOM | 925 | CA | GLU A 129 | 79.189 | 10.629 | 19.925 | 1.00 | 22.36 | A | C |
| ATOM | 926 | CB | GLU A 129 | 80.581 | 11.202 | 20.226 | 1.00 | 22.99 | A | C |
| ATOM | 927 | CG | GLU A 129 | 80.876 | 11.450 | 21.713 | 1.00 | 22.81 | A | C |
| ATOM | 928 | CD | GLU A 129 | 80.434 | 12.811 | 22.212 | 1.00 | 24.01 | A | C |
| ATOM | 929 | OE1 | GLU A 129 | 79.379 | 13.304 | 21.760 | 1.00 | 23.48 | A | O |
| ATOM | 930 | OE2 | GLU A 129 | 81.136 | 13.382 | 23.076 | 1.00 | 23.22 | A | O |
| ATOM | 931 | C | GLU A 129 | 78.829 | 9.594 | 20.996 | 1.00 | 23.83 | A | C |
| ATOM | 932 | O | GLU A 129 | 78.200 | 9.917 | 22.003 | 1.00 | 24.41 | A | O |
| ATOM | 933 | N | THR A 130 | 79.242 | 8.350 | 20.777 | 1.00 | 24.74 | A | N |
| ATOM | 934 | CA | THR A 130 | 78.980 | 7.270 | 21.726 | 1.00 | 26.14 | A | C |
| ATOM | 935 | CB | THR A 130 | 79.121 | 5.883 | 21.043 | 1.00 | 27.45 | A | C |
| ATOM | 936 | OG1 | THR A 130 | 80.456 | 5.714 | 20.537 | 1.00 | 26.95 | A | O |
| ATOM | 937 | CG2 | THR A 130 | 78.129 | 5.766 | 19.879 | 1.00 | 27.81 | A | C |
| ATOM | 938 | C | THR A 130 | 79.935 | 7.336 | 22.919 | 1.00 | 27.21 | A | C |
| ATOM | 939 | O | THR A 130 | 80.861 | 8.144 | 22.941 | 1.00 | 27.76 | A | O |
| ATOM | 940 | N | ARG A 131 | 79.695 | 6.504 | 23.922 | 1.00 | 28.47 | A | N |
| ATOM | 941 | CA | ARG A 131 | 80.557 | 6.485 | 25.085 | 1.00 | 30.46 | A | C |
| ATOM | 942 | CB | ARG A 131 | 80.114 | 5.421 | 26.079 | 1.00 | 32.07 | A | C |
| ATOM | 943 | CG | ARG A 131 | 81.134 | 5.238 | 27.175 | 1.00 | 36.56 | A | C |
| ATOM | 944 | CD | ARG A 131 | 80.673 | 4.298 | 28.256 | 1.00 | 40.21 | A | C |
| ATOM | 945 | NE | ARG A 131 | 81.463 | 4.521 | 29.466 | 1.00 | 43.27 | A | N |
| ATOM | 946 | CZ | ARG A 131 | 81.258 | 3.899 | 30.622 | 1.00 | 44.93 | A | C |
| ATOM | 947 | NH1 | ARG A 131 | 80.281 | 2.998 | 30.737 | 1.00 | 45.87 | A | N |
| ATOM | 948 | NH2 | ARG A 131 | 82.022 | 4.192 | 31.667 | 1.00 | 46.87 | A | N |
| ATOM | 949 | C | ARG A 131 | 81.995 | 6.197 | 24.681 | 1.00 | 31.02 | A | C |
| ATOM | 950 | O | ARG A 131 | 82.913 | 6.922 | 25.063 | 1.00 | 32.25 | A | O |
| ATOM | 951 | N | ALA A 132 | 82.186 | 5.134 | 23.907 | 1.00 | 29.92 | A | N |
| ATOM | 952 | CA | ALA A 132 | 83.512 | 4.750 | 23.461 | 1.00 | 29.99 | A | C |
| ATOM | 953 | CB | ALA A 132 | 83.438 | 3.449 | 22.661 | 1.00 | 31.07 | A | C |
| ATOM | 954 | C | ALA A 132 | 84.193 | 5.833 | 22.635 | 1.00 | 30.21 | A | C |
| ATOM | 955 | O | ALA A 132 | 85.367 | 6.121 | 22.844 | 1.00 | 30.04 | A | O |
| ATOM | 956 | N | GLU A 133 | 83.466 | 6.434 | 21.693 | 1.00 | 30.43 | A | N |
| ATOM | 957 | CA | GLU A 133 | 84.053 | 7.482 | 20.857 | 1.00 | 29.12 | A | C |
| ATOM | 958 | CB | GLU A 133 | 83.024 | 7.982 | 19.834 | 1.00 | 30.35 | A | C |
| ATOM | 959 | CG | GLU A 133 | 82.484 | 6.885 | 18.911 | 1.00 | 29.99 | A | C |
| ATOM | 960 | CD | GLU A 133 | 81.501 | 7.400 | 17.866 | 1.00 | 30.52 | A | C |
| ATOM | 961 | OE1 | GLU A 133 | 80.541 | 8.112 | 18.226 | 1.00 | 29.51 | A | O |
| ATOM | 962 | OE2 | GLU A 133 | 81.682 | 7.080 | 16.675 | 1.00 | 30.64 | A | O |
| ATOM | 963 | C | GLU A 133 | 84.580 | 8.641 | 21.718 | 1.00 | 29.17 | A | C |
| ATOM | 964 | O | GLU A 133 | 85.686 | 9.142 | 21.496 | 1.00 | 26.87 | A | O |
| ATOM | 965 | N | ARG A 134 | 83.797 | 9.049 | 22.712 | 1.00 | 27.71 | A | N |
| ATOM | 966 | CA | ARG A 134 | 84.202 | 10.129 | 23.606 | 1.00 | 28.85 | A | C |
| ATOM | 967 | CB | ARG A 134 | 83.059 | 10.493 | 24.555 | 1.00 | 28.22 | A | C |
| ATOM | 968 | CG | ARG A 134 | 83.249 | 11.810 | 25.317 | 1.00 | 29.74 | A | C |
| ATOM | 969 | CD | ARG A 134 | 82.127 | 12.030 | 26.338 | 1.00 | 28.01 | A | C |
| ATOM | 970 | NE | ARG A 134 | 82.295 | 13.255 | 27.116 | 1.00 | 27.65 | A | N |
| ATOM | 971 | CZ | ARG A 134 | 82.032 | 14.478 | 26.667 | 1.00 | 25.70 | A | C |
| ATOM | 972 | NH1 | ARG A 134 | 81.577 | 14.657 | 25.433 | 1.00 | 24.77 | A | N |
| ATOM | 973 | NH2 | ARG A 134 | 82.235 | 15.525 | 27.456 | 1.00 | 26.29 | A | N |
| ATOM | 974 | C | ARG A 134 | 85.433 | 9.713 | 24.418 | 1.00 | 30.81 | A | C |
| ATOM | 975 | O | ARG A 134 | 86.396 | 10.473 | 24.529 | 1.00 | 30.27 | A | O |
| ATOM | 976 | N | GLU A 135 | 85.399 | 8.511 | 24.991 | 1.00 | 32.23 | A | N |
| ATOM | 977 | CA | GLU A 135 | 86.530 | 8.017 | 25.767 | 1.00 | 32.91 | A | C |
| ATOM | 978 | CB | GLU A 135 | 86.220 | 6.633 | 26.335 | 1.00 | 35.01 | A | C |
| ATOM | 979 | CG | GLU A 135 | 85.423 | 6.684 | 27.638 | 1.00 | 37.98 | A | C |
| ATOM | 980 | CD | GLU A 135 | 84.703 | 5.378 | 27.955 | 1.00 | 39.91 | A | C |
| ATOM | 981 | OE1 | GLU A 135 | 84.012 | 5.318 | 28.997 | 1.00 | 41.05 | A | O |
| ATOM | 982 | OE2 | GLU A 135 | 84.813 | 4.418 | 27.164 | 1.00 | 40.24 | A | O |
| ATOM | 983 | C | GLU A 135 | 87.748 | 7.956 | 24.862 | 1.00 | 32.70 | A | C |
| ATOM | 984 | O | GLU A 135 | 88.859 | 8.278 | 25.278 | 1.00 | 32.55 | A | O |
| ATOM | 985 | N | ALA A 136 | 87.531 | 7.544 | 23.618 | 1.00 | 31.91 | A | N |
| ATOM | 986 | CA | ALA A 136 | 88.610 | 7.465 | 22.639 | 1.00 | 31.75 | A | C |
| ATOM | 987 | CB | ALA A 136 | 88.149 | 6.681 | 21.407 | 1.00 | 28.55 | A | C |

FIGURE 9- 15

SUBSTITUTE SHEET (RULE 26)

25/76

| | | | | | | | | | | |
|------|------|-----|-----------|--------|--------|--------|------|-------|---|---|
| ATOM | 988 | C | ALA A 136 | 89.050 | 8.881 | 22.241 | 1.00 | 31.53 | A | C |
| ATOM | 989 | O | ALA A 136 | 89.965 | 9.061 | 21.442 | 1.00 | 32.95 | A | O |
| ATOM | 990 | N | GLY A 137 | 88.387 | 9.878 | 22.818 | 1.00 | 32.61 | A | N |
| ATOM | 991 | CA | GLY A 137 | 88.704 | 11.270 | 22.539 | 1.00 | 31.69 | A | C |
| ATOM | 992 | C | GLY A 137 | 88.410 | 11.728 | 21.125 | 1.00 | 31.34 | A | C |
| ATOM | 993 | O | GLY A 137 | 89.127 | 12.567 | 20.586 | 1.00 | 31.03 | A | O |
| ATOM | 994 | N | LYS A 138 | 87.340 | 11.209 | 20.531 | 1.00 | 31.03 | A | N |
| ATOM | 995 | CA | LYS A 138 | 87.001 | 11.564 | 19.158 | 1.00 | 30.58 | A | C |
| ATOM | 996 | CB | LYS A 138 | 86.906 | 10.281 | 18.315 | 1.00 | 32.89 | A | C |
| ATOM | 997 | CG | LYS A 138 | 88.175 | 9.416 | 18.426 | 1.00 | 35.67 | A | C |
| ATOM | 998 | CD | LYS A 138 | 88.354 | 8.382 | 17.294 | 1.00 | 38.93 | A | C |
| ATOM | 999 | CE | LYS A 138 | 87.308 | 7.249 | 17.322 | 1.00 | 40.48 | A | C |
| ATOM | 1000 | NZ | LYS A 138 | 86.009 | 7.617 | 16.656 | 1.00 | 40.96 | A | N |
| ATOM | 1001 | C | LYS A 138 | 85.739 | 12.403 | 18.993 | 1.00 | 29.05 | A | C |
| ATOM | 1002 | O | LYS A 138 | 85.197 | 12.507 | 17.888 | 1.00 | 28.84 | A | O |
| ATOM | 1003 | N | THR A 139 | 85.284 | 13.024 | 20.080 | 1.00 | 28.96 | A | N |
| ATOM | 1004 | CA | THR A 139 | 84.074 | 13.846 | 20.038 | 1.00 | 26.71 | A | C |
| ATOM | 1005 | CB | THR A 139 | 83.917 | 14.684 | 21.314 | 1.00 | 26.35 | A | C |
| ATOM | 1006 | OG1 | THR A 139 | 83.597 | 13.821 | 22.414 | 1.00 | 26.63 | A | O |
| ATOM | 1007 | CG2 | THR A 139 | 82.802 | 15.715 | 21.142 | 1.00 | 24.81 | A | C |
| ATOM | 1008 | C | THR A 139 | 84.035 | 14.790 | 18.842 | 1.00 | 27.01 | A | C |
| ATOM | 1009 | O | THR A 139 | 83.104 | 14.742 | 18.029 | 1.00 | 27.22 | A | O |
| ATOM | 1010 | N | LEU A 140 | 85.043 | 15.647 | 18.738 | 1.00 | 25.70 | A | N |
| ATOM | 1011 | CA | LEU A 140 | 85.110 | 16.598 | 17.644 | 1.00 | 26.48 | A | C |
| ATOM | 1012 | CB | LEU A 140 | 86.309 | 17.539 | 17.828 | 1.00 | 27.47 | A | C |
| ATOM | 1013 | CG | LEU A 140 | 86.282 | 18.361 | 19.125 | 1.00 | 29.93 | A | C |
| ATOM | 1014 | CD1 | LEU A 140 | 87.355 | 19.426 | 19.085 | 1.00 | 28.77 | A | C |
| ATOM | 1015 | CD2 | LEU A 140 | 84.920 | 19.012 | 19.286 | 1.00 | 30.12 | A | C |
| ATOM | 1016 | C | LEU A 140 | 85.182 | 15.909 | 16.287 | 1.00 | 25.39 | A | C |
| ATOM | 1017 | O | LEU A 140 | 84.527 | 16.339 | 15.343 | 1.00 | 24.49 | A | O |
| ATOM | 1018 | N | GLU A 141 | 85.975 | 14.845 | 16.189 | 1.00 | 25.50 | A | N |
| ATOM | 1019 | CA | GLU A 141 | 86.109 | 14.111 | 14.932 | 1.00 | 26.95 | A | C |
| ATOM | 1020 | CB | GLU A 141 | 87.073 | 12.927 | 15.071 | 1.00 | 27.24 | A | C |
| ATOM | 1021 | CG | GLU A 141 | 87.254 | 12.166 | 13.762 | 1.00 | 32.14 | A | C |
| ATOM | 1022 | CD | GLU A 141 | 87.917 | 10.806 | 13.934 | 1.00 | 35.40 | A | C |
| ATOM | 1023 | OE1 | GLU A 141 | 89.020 | 10.747 | 14.524 | 1.00 | 36.45 | A | O |
| ATOM | 1024 | OE2 | GLU A 141 | 87.335 | 9.794 | 13.471 | 1.00 | 36.32 | A | O |
| ATOM | 1025 | C | GLU A 141 | 84.763 | 13.573 | 14.479 | 1.00 | 26.89 | A | C |
| ATOM | 1026 | O | GLU A 141 | 84.409 | 13.659 | 13.304 | 1.00 | 27.94 | A | O |
| ATOM | 1027 | N | VAL A 142 | 84.023 | 13.002 | 15.419 | 1.00 | 27.40 | A | N |
| ATOM | 1028 | CA | VAL A 142 | 82.718 | 12.439 | 15.115 | 1.00 | 26.75 | A | C |
| ATOM | 1029 | CB | VAL A 142 | 82.143 | 11.686 | 16.338 | 1.00 | 27.33 | A | C |
| ATOM | 1030 | CG1 | VAL A 142 | 80.716 | 11.236 | 16.062 | 1.00 | 27.37 | A | C |
| ATOM | 1031 | CG2 | VAL A 142 | 83.015 | 10.465 | 16.651 | 1.00 | 27.45 | A | C |
| ATOM | 1032 | C | VAL A 142 | 81.735 | 13.509 | 14.653 | 1.00 | 26.36 | A | C |
| ATOM | 1033 | O | VAL A 142 | 81.099 | 13.344 | 13.620 | 1.00 | 25.50 | A | O |
| ATOM | 1034 | N | VAL A 143 | 81.610 | 14.605 | 15.398 | 1.00 | 25.40 | A | N |
| ATOM | 1035 | CA | VAL A 143 | 80.677 | 15.657 | 14.999 | 1.00 | 26.36 | A | C |
| ATOM | 1036 | CB | VAL A 143 | 80.532 | 16.767 | 16.087 | 1.00 | 27.76 | A | C |
| ATOM | 1037 | CG1 | VAL A 143 | 79.939 | 16.173 | 17.355 | 1.00 | 29.26 | A | C |
| ATOM | 1038 | CG2 | VAL A 143 | 81.878 | 17.416 | 16.372 | 1.00 | 28.69 | A | C |
| ATOM | 1039 | C | VAL A 143 | 81.087 | 16.293 | 13.667 | 1.00 | 26.94 | A | C |
| ATOM | 1040 | O | VAL A 143 | 80.232 | 16.681 | 12.874 | 1.00 | 27.57 | A | O |
| ATOM | 1041 | N | ALA A 144 | 82.389 | 16.396 | 13.418 | 1.00 | 25.74 | A | N |
| ATOM | 1042 | CA | ALA A 144 | 82.867 | 16.953 | 12.157 | 1.00 | 27.44 | A | C |
| ATOM | 1043 | CB | ALA A 144 | 84.388 | 17.144 | 12.200 | 1.00 | 27.27 | A | C |
| ATOM | 1044 | C | ALA A 144 | 82.490 | 16.013 | 11.012 | 1.00 | 27.04 | A | C |
| ATOM | 1045 | O | ALA A 144 | 82.268 | 16.442 | 9.882 | 1.00 | 29.25 | A | O |
| ATOM | 1046 | N | ARG A 145 | 82.426 | 14.724 | 11.301 | 1.00 | 27.56 | A | N |
| ATOM | 1047 | CA | ARG A 145 | 82.066 | 13.752 | 10.279 | 1.00 | 27.25 | A | C |
| ATOM | 1048 | CB | ARG A 145 | 82.334 | 12.330 | 10.791 | 1.00 | 26.96 | A | C |
| ATOM | 1049 | CG | ARG A 145 | 82.170 | 11.216 | 9.752 | 1.00 | 30.84 | A | C |
| ATOM | 1050 | CD | ARG A 145 | 82.402 | 9.843 | 10.383 | 1.00 | 30.88 | A | C |
| ATOM | 1051 | NE | ARG A 145 | 81.405 | 9.547 | 11.411 | 1.00 | 34.11 | A | N |
| ATOM | 1052 | CZ | ARG A 145 | 81.687 | 9.109 | 12.637 | 1.00 | 36.26 | A | C |
| ATOM | 1053 | NH1 | ARG A 145 | 82.945 | 8.906 | 13.019 | 1.00 | 37.82 | A | N |

FIGURE 9- 16

SUBSTITUTE SHEET (RULE 26)

26/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 1054 | NH2 | ARG | A | 145 | 80.702 | 8.870 | 13.493 | 1.00 | 40.90 | A | N |
| ATOM | 1055 | C | ARG | A | 145 | 80.584 | 13.918 | 9.927 | 1.00 | 27.29 | A | C |
| ATOM | 1056 | O | ARG | A | 145 | 80.230 | 14.138 | 8.771 | 1.00 | 28.69 | A | O |
| ATOM | 1057 | N | GLN | A | 146 | 79.726 | 13.845 | 10.939 | 1.00 | 27.04 | A | N |
| ATOM | 1058 | CA | GLN | A | 146 | 78.281 | 13.943 | 10.738 | 1.00 | 25.27 | A | C |
| ATOM | 1059 | CB | GLN | A | 146 | 77.570 | 13.679 | 12.068 | 1.00 | 24.31 | A | C |
| ATOM | 1060 | CG | GLN | A | 146 | 77.959 | 12.334 | 12.664 | 1.00 | 24.79 | A | C |
| ATOM | 1061 | CD | GLN | A | 146 | 77.352 | 12.066 | 14.037 | 1.00 | 26.11 | A | C |
| ATOM | 1062 | OE1 | GLN | A | 146 | 77.351 | 12.935 | 14.916 | 1.00 | 23.99 | A | O |
| ATOM | 1063 | NE2 | GLN | A | 146 | 76.854 | 10.844 | 14.232 | 1.00 | 23.71 | A | N |
| ATOM | 1064 | C | GLN | A | 146 | 77.817 | 15.259 | 10.118 | 1.00 | 25.40 | A | C |
| ATOM | 1065 | O | GLN | A | 146 | 76.833 | 15.289 | 9.378 | 1.00 | 25.59 | A | O |
| ATOM | 1066 | N | LEU | A | 147 | 78.519 | 16.343 | 10.422 | 1.00 | 24.90 | A | N |
| ATOM | 1067 | CA | LEU | A | 147 | 78.186 | 17.650 | 9.859 | 1.00 | 25.07 | A | C |
| ATOM | 1068 | CB | LEU | A | 147 | 78.852 | 18.764 | 10.665 | 1.00 | 22.74 | A | C |
| ATOM | 1069 | CG | LEU | A | 147 | 78.400 | 18.987 | 12.104 | 1.00 | 21.74 | A | C |
| ATOM | 1070 | CD1 | LEU | A | 147 | 79.322 | 19.987 | 12.758 | 1.00 | 22.89 | A | C |
| ATOM | 1071 | CD2 | LEU | A | 147 | 76.967 | 19.487 | 12.122 | 1.00 | 24.07 | A | C |
| ATOM | 1072 | C | LEU | A | 147 | 78.703 | 17.707 | 8.423 | 1.00 | 25.67 | A | C |
| ATOM | 1073 | O | LEU | A | 147 | 78.092 | 18.321 | 7.544 | 1.00 | 25.46 | A | O |
| ATOM | 1074 | N | GLY | A | 148 | 79.839 | 17.051 | 8.203 | 1.00 | 27.56 | A | N |
| ATOM | 1075 | CA | GLY | A | 148 | 80.459 | 17.032 | 6.890 | 1.00 | 29.17 | A | C |
| ATOM | 1076 | C | GLY | A | 148 | 79.660 | 16.381 | 5.777 | 1.00 | 30.48 | A | C |
| ATOM | 1077 | O | GLY | A | 148 | 79.768 | 16.800 | 4.617 | 1.00 | 30.01 | A | O |
| ATOM | 1078 | N | SER | A | 149 | 78.863 | 15.365 | 6.098 | 1.00 | 29.94 | A | N |
| ATOM | 1079 | CA | SER | A | 149 | 78.097 | 14.714 | 5.050 | 1.00 | 32.34 | A | C |
| ATOM | 1080 | CB | SER | A | 149 | 77.250 | 13.555 | 5.607 | 1.00 | 35.10 | A | C |
| ATOM | 1081 | OG | SER | A | 149 | 76.235 | 14.006 | 6.492 | 1.00 | 37.50 | A | O |
| ATOM | 1082 | C | SER | A | 149 | 77.215 | 15.757 | 4.392 | 1.00 | 31.10 | A | C |
| ATOM | 1083 | O | SER | A | 149 | 77.166 | 15.867 | 3.166 | 1.00 | 33.27 | A | O |
| ATOM | 1084 | N | VAL | A | 150 | 76.532 | 16.545 | 5.207 | 1.00 | 29.54 | A | N |
| ATOM | 1085 | CA | VAL | A | 150 | 75.668 | 17.584 | 4.673 | 1.00 | 28.08 | A | C |
| ATOM | 1086 | CB | VAL | A | 150 | 74.891 | 18.259 | 5.796 | 1.00 | 27.41 | A | C |
| ATOM | 1087 | CG1 | VAL | A | 150 | 74.033 | 19.381 | 5.242 | 1.00 | 28.70 | A | C |
| ATOM | 1088 | CG2 | VAL | A | 150 | 74.031 | 17.226 | 6.494 | 1.00 | 30.30 | A | C |
| ATOM | 1089 | C | VAL | A | 150 | 76.477 | 18.633 | 3.908 | 1.00 | 27.76 | A | C |
| ATOM | 1090 | O | VAL | A | 150 | 76.092 | 19.049 | 2.817 | 1.00 | 26.92 | A | O |
| ATOM | 1091 | N | ILE | A | 151 | 77.598 | 19.058 | 4.477 | 1.00 | 27.72 | A | N |
| ATOM | 1092 | CA | ILE | A | 151 | 78.435 | 20.065 | 3.824 | 1.00 | 28.74 | A | C |
| ATOM | 1093 | CB | ILE | A | 151 | 79.674 | 20.414 | 4.699 | 1.00 | 28.30 | A | C |
| ATOM | 1094 | CG2 | ILE | A | 151 | 80.762 | 21.059 | 3.842 | 1.00 | 27.06 | A | C |
| ATOM | 1095 | CG1 | ILE | A | 151 | 79.243 | 21.321 | 5.859 | 1.00 | 29.37 | A | C |
| ATOM | 1096 | CD1 | ILE | A | 151 | 80.353 | 21.619 | 6.879 | 1.00 | 29.91 | A | C |
| ATOM | 1097 | C | ILE | A | 151 | 78.913 | 19.628 | 2.440 | 1.00 | 29.17 | A | C |
| ATOM | 1098 | O | ILE | A | 151 | 78.874 | 20.410 | 1.486 | 1.00 | 28.72 | A | O |
| ATOM | 1099 | N | ASP | A | 152 | 79.370 | 18.382 | 2.332 | 1.00 | 29.50 | A | N |
| ATOM | 1100 | CA | ASP | A | 152 | 79.856 | 17.863 | 1.058 | 1.00 | 30.02 | A | C |
| ATOM | 1101 | CB | ASP | A | 152 | 80.384 | 16.432 | 1.231 | 1.00 | 32.67 | A | C |
| ATOM | 1102 | CG | ASP | A | 152 | 81.705 | 16.381 | 1.981 | 1.00 | 36.20 | A | C |
| ATOM | 1103 | OD1 | ASP | A | 152 | 82.707 | 16.903 | 1.442 | 1.00 | 39.62 | A | O |
| ATOM | 1104 | OD2 | ASP | A | 152 | 81.747 | 15.823 | 3.102 | 1.00 | 38.19 | A | O |
| ATOM | 1105 | C | ASP | A | 152 | 78.783 | 17.878 | -0.025 | 1.00 | 28.90 | A | C |
| ATOM | 1106 | O | ASP | A | 152 | 79.093 | 18.007 | -1.207 | 1.00 | 28.85 | A | O |
| ATOM | 1107 | N | GLU | A | 153 | 77.523 | 17.756 | 0.373 | 1.00 | 28.78 | A | N |
| ATOM | 1108 | CA | GLU | A | 153 | 76.423 | 17.737 | -0.596 | 1.00 | 29.97 | A | C |
| ATOM | 1109 | CB | GLU | A | 153 | 75.273 | 16.853 | -0.094 | 1.00 | 32.44 | A | C |
| ATOM | 1110 | CG | GLU | A | 153 | 75.682 | 15.427 | 0.210 | 1.00 | 38.67 | A | C |
| ATOM | 1111 | CD | GLU | A | 153 | 76.345 | 14.757 | -0.977 | 1.00 | 40.99 | A | C |
| ATOM | 1112 | OE1 | GLU | A | 153 | 77.273 | 13.949 | -0.728 | 1.00 | 43.21 | A | O |
| ATOM | 1113 | OE2 | GLU | A | 153 | 75.940 | 15.034 | -2.141 | 1.00 | 39.50 | A | O |
| ATOM | 1114 | C | GLU | A | 153 | 75.830 | 19.088 | -0.960 | 1.00 | 28.33 | A | C |
| ATOM | 1115 | O | GLU | A | 153 | 75.524 | 19.334 | -2.125 | 1.00 | 27.07 | A | O |
| ATOM | 1116 | N | LEU | A | 154 | 75.654 | 19.950 | 0.035 | 1.00 | 25.22 | A | N |
| ATOM | 1117 | CA | LEU | A | 154 | 75.028 | 21.248 | -0.189 | 1.00 | 25.00 | A | C |
| ATOM | 1118 | CB | LEU | A | 154 | 73.701 | 21.311 | 0.569 | 1.00 | 24.29 | A | C |
| ATOM | 1119 | CG | LEU | A | 154 | 72.739 | 20.121 | 0.496 | 1.00 | 25.26 | A | C |

FIGURE 9 - 17

SUBSTITUTE SHEET (RULE 26)

27/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 1120 | CD1 | LEU | A | 154 | 71.677 | 20.301 | 1.579 | 1.00 | 25.02 | A | C |
| ATOM | 1121 | CD2 | LEU | A | 154 | 72.103 | 20.020 | -0.880 | 1.00 | 25.41 | A | C |
| ATOM | 1122 | C | LEU | A | 154 | 75.857 | 22.456 | 0.220 | 1.00 | 24.07 | A | C |
| ATOM | 1123 | O | LEU | A | 154 | 75.454 | 23.588 | -0.032 | 1.00 | 26.23 | A | O |
| ATOM | 1124 | N | GLY | A | 155 | 76.998 | 22.225 | 0.854 | 1.00 | 23.88 | A | N |
| ATOM | 1125 | CA | GLY | A | 155 | 77.829 | 23.333 | 1.301 | 1.00 | 24.05 | A | C |
| ATOM | 1126 | C | GLY | A | 155 | 77.322 | 23.827 | 2.645 | 1.00 | 23.55 | A | C |
| ATOM | 1127 | O | GLY | A | 155 | 76.157 | 23.630 | 2.983 | 1.00 | 23.96 | A | O |
| ATOM | 1128 | N | VAL | A | 156 | 78.178 | 24.486 | 3.410 | 1.00 | 22.67 | A | N |
| ATOM | 1129 | CA | VAL | A | 156 | 77.788 | 24.962 | 4.727 | 1.00 | 23.92 | A | C |
| ATOM | 1130 | CB | VAL | A | 156 | 78.981 | 25.628 | 5.447 | 1.00 | 22.95 | A | C |
| ATOM | 1131 | CG1 | VAL | A | 156 | 79.373 | 26.902 | 4.725 | 1.00 | 23.73 | A | C |
| ATOM | 1132 | CG2 | VAL | A | 156 | 78.623 | 25.902 | 6.901 | 1.00 | 21.73 | A | C |
| ATOM | 1133 | C | VAL | A | 156 | 76.606 | 25.929 | 4.701 | 1.00 | 23.73 | A | C |
| ATOM | 1134 | O | VAL | A | 156 | 75.936 | 26.116 | 5.709 | 1.00 | 24.38 | A | O |
| ATOM | 1135 | N | GLY | A | 157 | 76.347 | 26.531 | 3.547 | 1.00 | 25.39 | A | N |
| ATOM | 1136 | CA | GLY | A | 157 | 75.239 | 27.463 | 3.431 | 1.00 | 26.59 | A | C |
| ATOM | 1137 | C | GLY | A | 157 | 73.891 | 26.822 | 3.717 | 1.00 | 28.69 | A | C |
| ATOM | 1138 | O | GLY | A | 157 | 72.940 | 27.505 | 4.107 | 1.00 | 29.93 | A | O |
| ATOM | 1139 | N | ALA | A | 158 | 73.805 | 25.508 | 3.537 | 1.00 | 27.95 | A | N |
| ATOM | 1140 | CA | ALA | A | 158 | 72.561 | 24.781 | 3.782 | 1.00 | 27.95 | A | C |
| ATOM | 1141 | CB | ALA | A | 158 | 72.738 | 23.297 | 3.423 | 1.00 | 26.06 | A | C |
| ATOM | 1142 | C | ALA | A | 158 | 72.100 | 24.918 | 5.237 | 1.00 | 26.98 | A | C |
| ATOM | 1143 | O | ALA | A | 158 | 70.908 | 24.839 | 5.528 | 1.00 | 25.53 | A | O |
| ATOM | 1144 | N | PHE | A | 159 | 73.049 | 25.128 | 6.142 | 1.00 | 25.75 | A | N |
| ATOM | 1145 | CA | PHE | A | 159 | 72.738 | 25.270 | 7.563 | 1.00 | 25.59 | A | C |
| ATOM | 1146 | CB | PHE | A | 159 | 74.019 | 25.157 | 8.387 | 1.00 | 24.12 | A | C |
| ATOM | 1147 | CG | PHE | A | 159 | 74.489 | 23.748 | 8.559 | 1.00 | 25.38 | A | C |
| ATOM | 1148 | CD1 | PHE | A | 159 | 73.818 | 22.882 | 9.415 | 1.00 | 26.14 | A | C |
| ATOM | 1149 | CD2 | PHE | A | 159 | 75.577 | 23.267 | 7.844 | 1.00 | 26.43 | A | C |
| ATOM | 1150 | CE1 | PHE | A | 159 | 74.221 | 21.560 | 9.552 | 1.00 | 25.07 | A | C |
| ATOM | 1151 | CE2 | PHE | A | 159 | 75.987 | 21.939 | 7.977 | 1.00 | 26.60 | A | C |
| ATOM | 1152 | CZ | PHE | A | 159 | 75.306 | 21.089 | 8.831 | 1.00 | 25.41 | A | C |
| ATOM | 1153 | C | PHE | A | 159 | 72.001 | 26.562 | 7.905 | 1.00 | 25.78 | A | C |
| ATOM | 1154 | O | PHE | A | 159 | 71.526 | 26.735 | 9.023 | 1.00 | 26.86 | A | O |
| ATOM | 1155 | N | ALA | A | 160 | 71.905 | 27.467 | 6.941 | 1.00 | 25.69 | A | N |
| ATOM | 1156 | CA | ALA | A | 160 | 71.199 | 28.717 | 7.160 | 1.00 | 26.51 | A | C |
| ATOM | 1157 | CB | ALA | A | 160 | 71.514 | 29.697 | 6.041 | 1.00 | 26.86 | A | C |
| ATOM | 1158 | C | ALA | A | 160 | 69.696 | 28.408 | 7.198 | 1.00 | 27.37 | A | C |
| ATOM | 1159 | O | ALA | A | 160 | 68.897 | 29.185 | 7.716 | 1.00 | 26.15 | A | O |
| ATOM | 1160 | N | ARG | A | 161 | 69.317 | 27.260 | 6.652 | 1.00 | 27.52 | A | N |
| ATOM | 1161 | CA | ARG | A | 161 | 67.916 | 26.863 | 6.655 | 1.00 | 28.68 | A | C |
| ATOM | 1162 | CB | ARG | A | 161 | 67.355 | 26.927 | 5.233 | 1.00 | 31.19 | A | C |
| ATOM | 1163 | CG | ARG | A | 161 | 66.920 | 28.343 | 4.845 | 1.00 | 34.54 | A | C |
| ATOM | 1164 | CD | ARG | A | 161 | 66.725 | 28.495 | 3.354 | 1.00 | 38.15 | A | C |
| ATOM | 1165 | NE | ARG | A | 161 | 67.992 | 28.735 | 2.662 | 1.00 | 44.59 | A | N |
| ATOM | 1166 | CZ | ARG | A | 161 | 68.532 | 29.941 | 2.465 | 1.00 | 46.31 | A | C |
| ATOM | 1167 | NH1 | ARG | A | 161 | 67.921 | 31.037 | 2.907 | 1.00 | 47.29 | A | N |
| ATOM | 1168 | NH2 | ARG | A | 161 | 69.681 | 30.054 | 1.808 | 1.00 | 48.04 | A | N |
| ATOM | 1169 | C | ARG | A | 161 | 67.740 | 25.481 | 7.276 | 1.00 | 26.99 | A | C |
| ATOM | 1170 | O | ARG | A | 161 | 67.032 | 24.618 | 6.755 | 1.00 | 25.86 | A | O |
| ATOM | 1171 | N | ALA | A | 162 | 68.387 | 25.306 | 8.423 | 1.00 | 26.87 | A | N |
| ATOM | 1172 | CA | ALA | A | 162 | 68.348 | 24.059 | 9.164 | 1.00 | 25.70 | A | C |
| ATOM | 1173 | CB | ALA | A | 162 | 69.356 | 23.092 | 8.593 | 1.00 | 25.54 | A | C |
| ATOM | 1174 | C | ALA | A | 162 | 68.650 | 24.309 | 10.634 | 1.00 | 26.90 | A | C |
| ATOM | 1175 | O | ALA | A | 162 | 68.913 | 25.443 | 11.043 | 1.00 | 27.80 | A | O |
| ATOM | 1176 | N | VAL | A | 163 | 68.611 | 23.237 | 11.417 | 1.00 | 26.24 | A | N |
| ATOM | 1177 | CA | VAL | A | 163 | 68.874 | 23.291 | 12.846 | 1.00 | 25.08 | A | C |
| ATOM | 1178 | CB | VAL | A | 163 | 67.559 | 23.221 | 13.666 | 1.00 | 25.69 | A | C |
| ATOM | 1179 | CG1 | VAL | A | 163 | 67.871 | 23.148 | 15.148 | 1.00 | 26.56 | A | C |
| ATOM | 1180 | CG2 | VAL | A | 163 | 66.689 | 24.438 | 13.372 | 1.00 | 26.92 | A | C |
| ATOM | 1181 | C | VAL | A | 163 | 69.721 | 22.080 | 13.197 | 1.00 | 23.50 | A | C |
| ATOM | 1182 | O | VAL | A | 163 | 69.608 | 21.043 | 12.565 | 1.00 | 24.62 | A | O |
| ATOM | 1183 | N | VAL | A | 164 | 70.586 | 22.211 | 14.189 | 1.00 | 23.33 | A | N |
| ATOM | 1184 | CA | VAL | A | 164 | 71.399 | 21.077 | 14.590 | 1.00 | 22.31 | A | C |
| ATOM | 1185 | CB | VAL | A | 164 | 72.899 | 21.361 | 14.469 | 1.00 | 20.94 | A | C |

FIGURE 9 - 18

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 1186 | CG1 | VAL | A | 164 | 73.683 | 20.173 | 14.987 | 1.00 | 20.70 | A | C |
| ATOM | 1187 | CG2 | VAL | A | 164 | 73.258 | 21.636 | 13.025 | 1.00 | 21.81 | A | C |
| ATOM | 1188 | C | VAL | A | 164 | 71.088 | 20.756 | 16.039 | 1.00 | 22.34 | A | C |
| ATOM | 1189 | O | VAL | A | 164 | 70.936 | 21.662 | 16.853 | 1.00 | 21.99 | A | O |
| ATOM | 1190 | N | ALA | A | 165 | 70.985 | 19.469 | 16.348 | 1.00 | 21.11 | A | N |
| ATOM | 1191 | CA | ALA | A | 165 | 70.708 | 19.027 | 17.712 | 1.00 | 20.36 | A | C |
| ATOM | 1192 | CB | ALA | A | 165 | 69.367 | 18.326 | 17.783 | 1.00 | 20.04 | A | C |
| ATOM | 1193 | C | ALA | A | 165 | 71.809 | 18.081 | 18.166 | 1.00 | 21.99 | A | C |
| ATOM | 1194 | O | ALA | A | 165 | 72.081 | 17.071 | 17.513 | 1.00 | 21.33 | A | O |
| ATOM | 1195 | N | TYR | A | 166 | 72.444 | 18.415 | 19.283 | 1.00 | 21.18 | A | N |
| ATOM | 1196 | CA | TYR | A | 166 | 73.506 | 17.580 | 19.817 | 1.00 | 21.78 | A | C |
| ATOM | 1197 | CB | TYR | A | 166 | 74.591 | 18.431 | 20.476 | 1.00 | 18.90 | A | C |
| ATOM | 1198 | CG | TYR | A | 166 | 75.585 | 17.595 | 21.246 | 1.00 | 18.81 | A | C |
| ATOM | 1199 | CD1 | TYR | A | 166 | 76.396 | 16.667 | 20.590 | 1.00 | 18.53 | A | C |
| ATOM | 1200 | CE1 | TYR | A | 166 | 77.288 | 15.866 | 21.292 | 1.00 | 17.92 | A | C |
| ATOM | 1201 | CD2 | TYR | A | 166 | 75.691 | 17.703 | 22.632 | 1.00 | 18.54 | A | C |
| ATOM | 1202 | CE2 | TYR | A | 166 | 76.586 | 16.903 | 23.351 | 1.00 | 19.94 | A | C |
| ATOM | 1203 | CZ | TYR | A | 166 | 77.381 | 15.988 | 22.672 | 1.00 | 20.01 | A | C |
| ATOM | 1204 | OH | TYR | A | 166 | 78.277 | 15.210 | 23.363 | 1.00 | 21.99 | A | O |
| ATOM | 1205 | C | TYR | A | 166 | 72.993 | 16.579 | 20.844 | 1.00 | 22.54 | A | C |
| ATOM | 1206 | O | TYR | A | 166 | 72.445 | 16.970 | 21.884 | 1.00 | 22.79 | A | O |
| ATOM | 1207 | N | GLU | A | 167 | 73.175 | 15.295 | 20.554 | 1.00 | 22.07 | A | N |
| ATOM | 1208 | CA | GLU | A | 167 | 72.760 | 14.245 | 21.482 | 1.00 | 23.13 | A | C |
| ATOM | 1209 | CB | GLU | A | 167 | 72.065 | 13.092 | 20.770 | 1.00 | 22.55 | A | C |
| ATOM | 1210 | CG | GLU | A | 167 | 70.927 | 13.425 | 19.866 | 1.00 | 25.06 | A | C |
| ATOM | 1211 | CD | GLU | A | 167 | 70.341 | 12.156 | 19.271 | 1.00 | 25.54 | A | C |
| ATOM | 1212 | OE1 | GLU | A | 167 | 69.662 | 11.415 | 20.012 | 1.00 | 26.54 | A | O |
| ATOM | 1213 | OE2 | GLU | A | 167 | 70.575 | 11.888 | 18.078 | 1.00 | 22.15 | A | O |
| ATOM | 1214 | C | GLU | A | 167 | 73.993 | 13.649 | 22.150 | 1.00 | 23.01 | A | C |
| ATOM | 1215 | O | GLU | A | 167 | 74.864 | 13.100 | 21.469 | 1.00 | 22.75 | A | O |
| ATOM | 1216 | N | PRO | A | 168 | 74.084 | 13.737 | 23.486 | 1.00 | 22.25 | A | N |
| ATOM | 1217 | CD | PRO | A | 168 | 73.207 | 14.432 | 24.446 | 1.00 | 22.41 | A | C |
| ATOM | 1218 | CA | PRO | A | 168 | 75.252 | 13.160 | 24.156 | 1.00 | 22.23 | A | C |
| ATOM | 1219 | CB | PRO | A | 168 | 75.286 | 13.910 | 25.482 | 1.00 | 22.75 | A | C |
| ATOM | 1220 | CG | PRO | A | 168 | 73.806 | 14.046 | 25.803 | 1.00 | 22.26 | A | C |
| ATOM | 1221 | C | PRO | A | 168 | 74.992 | 11.655 | 24.331 | 1.00 | 23.30 | A | C |
| ATOM | 1222 | O | PRO | A | 168 | 74.706 | 11.183 | 25.436 | 1.00 | 22.27 | A | O |
| ATOM | 1223 | N | VAL | A | 169 | 75.079 | 10.910 | 23.229 | 1.00 | 24.51 | A | N |
| ATOM | 1224 | CA | VAL | A | 169 | 74.829 | 9.469 | 23.253 | 1.00 | 24.97 | A | C |
| ATOM | 1225 | CB | VAL | A | 169 | 75.091 | 8.838 | 21.845 | 1.00 | 26.37 | A | C |
| ATOM | 1226 | CG1 | VAL | A | 169 | 74.927 | 7.311 | 21.899 | 1.00 | 27.34 | A | C |
| ATOM | 1227 | CG2 | VAL | A | 169 | 74.111 | 9.415 | 20.829 | 1.00 | 23.12 | A | C |
| ATOM | 1228 | C | VAL | A | 169 | 75.676 | 8.790 | 24.336 | 1.00 | 25.20 | A | C |
| ATOM | 1229 | O | VAL | A | 169 | 75.271 | 7.787 | 24.922 | 1.00 | 23.22 | A | O |
| ATOM | 1230 | N | TRP | A | 170 | 76.842 | 9.364 | 24.610 | 1.00 | 25.13 | A | N |
| ATOM | 1231 | CA | TRP | A | 170 | 77.741 | 8.850 | 25.635 | 1.00 | 25.43 | A | C |
| ATOM | 1232 | CB | TRP | A | 170 | 79.026 | 9.678 | 25.674 | 1.00 | 26.14 | A | C |
| ATOM | 1233 | CG | TRP | A | 170 | 78.810 | 11.181 | 25.802 | 1.00 | 25.69 | A | C |
| ATOM | 1234 | CD2 | TRP | A | 170 | 78.905 | 11.973 | 27.000 | 1.00 | 25.72 | A | C |
| ATOM | 1235 | CE2 | TRP | A | 170 | 78.727 | 13.327 | 26.624 | 1.00 | 24.81 | A | C |
| ATOM | 1236 | CE3 | TRP | A | 170 | 79.127 | 11.672 | 28.349 | 1.00 | 25.81 | A | C |
| ATOM | 1237 | CD1 | TRP | A | 170 | 78.571 | 12.065 | 24.785 | 1.00 | 26.18 | A | C |
| ATOM | 1238 | NE1 | TRP | A | 170 | 78.527 | 13.355 | 25.270 | 1.00 | 24.94 | A | N |
| ATOM | 1239 | CZ2 | TRP | A | 170 | 78.764 | 14.374 | 27.548 | 1.00 | 25.18 | A | C |
| ATOM | 1240 | CZ3 | TRP | A | 170 | 79.166 | 12.716 | 29.269 | 1.00 | 26.12 | A | C |
| ATOM | 1241 | CH2 | TRP | A | 170 | 78.985 | 14.051 | 28.862 | 1.00 | 25.30 | A | C |
| ATOM | 1242 | C | TRP | A | 170 | 77.091 | 8.903 | 27.015 | 1.00 | 26.95 | A | C |
| ATOM | 1243 | O | TRP | A | 170 | 77.519 | 8.212 | 27.937 | 1.00 | 26.86 | A | O |
| ATOM | 1244 | N | ALA | A | 171 | 76.063 | 9.733 | 27.155 | 1.00 | 28.53 | A | N |
| ATOM | 1245 | CA | ALA | A | 171 | 75.366 | 9.888 | 28.424 | 1.00 | 31.53 | A | C |
| ATOM | 1246 | CB | ALA | A | 171 | 75.455 | 11.331 | 28.896 | 1.00 | 30.64 | A | C |
| ATOM | 1247 | C | ALA | A | 171 | 73.911 | 9.487 | 28.296 | 1.00 | 34.29 | A | C |
| ATOM | 1248 | O | ALA | A | 171 | 73.084 | 9.840 | 29.135 | 1.00 | 35.05 | A | O |
| ATOM | 1249 | N | ILE | A | 172 | 73.584 | 8.757 | 27.237 | 1.00 | 36.78 | A | N |
| ATOM | 1250 | CA | ILE | A | 172 | 72.207 | 8.342 | 27.054 | 1.00 | 38.15 | A | C |
| ATOM | 1251 | CB | ILE | A | 172 | 71.670 | 8.772 | 25.656 | 1.00 | 38.81 | A | C |

FIGURE 9 - 19

29/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 1252 | CG2 | ILE | A | 172 | 70.229 | 8.294 | 25.468 | 1.00 | 39.21 | A | C |
| ATOM | 1253 | CG1 | ILE | A | 172 | 71.696 | 10.304 | 25.542 | 1.00 | 38.50 | A | C |
| ATOM | 1254 | CD1 | ILE | A | 172 | 71.036 | 10.850 | 24.273 | 1.00 | 38.17 | A | C |
| ATOM | 1255 | C | ILE | A | 172 | 72.082 | 6.843 | 27.263 | 1.00 | 39.10 | A | C |
| ATOM | 1256 | O | ILE | A | 172 | 72.300 | 6.046 | 26.350 | 1.00 | 39.16 | A | O |
| ATOM | 1257 | N | GLY | A | 173 | 71.756 | 6.482 | 28.502 | 1.00 | 41.09 | A | N |
| ATOM | 1258 | CA | GLY | A | 173 | 71.572 | 5.094 | 28.882 | 1.00 | 42.26 | A | C |
| ATOM | 1259 | C | GLY | A | 173 | 72.851 | 4.369 | 29.228 | 1.00 | 42.67 | A | C |
| ATOM | 1260 | O | GLY | A | 173 | 72.910 | 3.141 | 29.147 | 1.00 | 44.10 | A | O |
| ATOM | 1261 | N | THR | A | 174 | 73.867 | 5.117 | 29.640 | 1.00 | 42.67 | A | N |
| ATOM | 1262 | CA | THR | A | 174 | 75.164 | 4.532 | 29.964 | 1.00 | 40.83 | A | C |
| ATOM | 1263 | CB | THR | A | 174 | 76.263 | 5.246 | 29.198 | 1.00 | 42.40 | A | C |
| ATOM | 1264 | OG1 | THR | A | 174 | 76.254 | 6.628 | 29.587 | 1.00 | 41.90 | A | O |
| ATOM | 1265 | CG2 | THR | A | 174 | 76.028 | 5.146 | 27.693 | 1.00 | 42.07 | A | C |
| ATOM | 1266 | C | THR | A | 174 | 75.491 | 4.708 | 31.432 | 1.00 | 40.25 | A | C |
| ATOM | 1267 | O | THR | A | 174 | 76.461 | 4.135 | 31.933 | 1.00 | 38.57 | A | O |
| ATOM | 1268 | N | GLY | A | 175 | 74.690 | 5.520 | 32.114 | 1.00 | 39.49 | A | N |
| ATOM | 1269 | CA | GLY | A | 175 | 74.955 | 5.791 | 33.512 | 1.00 | 40.13 | A | C |
| ATOM | 1270 | C | GLY | A | 175 | 75.710 | 7.107 | 33.660 | 1.00 | 39.39 | A | C |
| ATOM | 1271 | O | GLY | A | 175 | 75.597 | 7.771 | 34.683 | 1.00 | 39.82 | A | O |
| ATOM | 1272 | N | LEU | A | 176 | 76.476 | 7.494 | 32.641 | 1.00 | 38.55 | A | N |
| ATOM | 1273 | CA | LEU | A | 176 | 77.226 | 8.752 | 32.692 | 1.00 | 36.30 | A | C |
| ATOM | 1274 | CB | LEU | A | 176 | 78.254 | 8.816 | 31.558 | 1.00 | 34.97 | A | C |
| ATOM | 1275 | CG | LEU | A | 176 | 79.221 | 7.625 | 31.496 | 1.00 | 34.55 | A | C |
| ATOM | 1276 | CD1 | LEU | A | 176 | 80.107 | 7.741 | 30.275 | 1.00 | 33.13 | A | C |
| ATOM | 1277 | CD2 | LEU | A | 176 | 80.054 | 7.565 | 32.768 | 1.00 | 33.82 | A | C |
| ATOM | 1278 | C | LEU | A | 176 | 76.240 | 9.901 | 32.557 | 1.00 | 36.22 | A | C |
| ATOM | 1279 | O | LEU | A | 176 | 75.171 | 9.739 | 31.962 | 1.00 | 36.88 | A | O |
| ATOM | 1280 | N | THR | A | 177 | 76.593 | 11.056 | 33.112 | 1.00 | 34.99 | A | N |
| ATOM | 1281 | CA | THR | A | 177 | 75.726 | 12.223 | 33.049 | 1.00 | 34.70 | A | C |
| ATOM | 1282 | CB | THR | A | 177 | 75.237 | 12.629 | 34.450 | 1.00 | 36.79 | A | C |
| ATOM | 1283 | OG1 | THR | A | 177 | 76.372 | 12.921 | 35.271 | 1.00 | 37.89 | A | O |
| ATOM | 1284 | CG2 | THR | A | 177 | 74.429 | 11.492 | 35.104 | 1.00 | 37.23 | A | C |
| ATOM | 1285 | C | THR | A | 177 | 76.439 | 13.427 | 32.440 | 1.00 | 33.78 | A | C |
| ATOM | 1286 | O | THR | A | 177 | 77.624 | 13.666 | 32.696 | 1.00 | 33.54 | A | O |
| ATOM | 1287 | N | ALA | A | 178 | 75.711 | 14.179 | 31.624 | 1.00 | 30.84 | A | N |
| ATOM | 1288 | CA | ALA | A | 178 | 76.266 | 15.365 | 31.010 | 1.00 | 28.77 | A | C |
| ATOM | 1289 | CB | ALA | A | 178 | 75.806 | 15.481 | 29.563 | 1.00 | 27.66 | A | C |
| ATOM | 1290 | C | ALA | A | 178 | 75.780 | 16.559 | 31.815 | 1.00 | 27.19 | A | C |
| ATOM | 1291 | O | ALA | A | 178 | 74.593 | 16.678 | 32.115 | 1.00 | 27.90 | A | O |
| ATOM | 1292 | N | SER | A | 179 | 76.709 | 17.430 | 32.179 | 1.00 | 25.65 | A | N |
| ATOM | 1293 | CA | SER | A | 179 | 76.393 | 18.629 | 32.936 | 1.00 | 24.89 | A | C |
| ATOM | 1294 | CB | SER | A | 179 | 77.538 | 18.958 | 33.891 | 1.00 | 25.65 | A | C |
| ATOM | 1295 | OG | SER | A | 179 | 78.713 | 19.255 | 33.157 | 1.00 | 26.02 | A | O |
| ATOM | 1296 | C | SER | A | 179 | 76.234 | 19.774 | 31.941 | 1.00 | 25.60 | A | C |
| ATOM | 1297 | O | SER | A | 179 | 76.673 | 19.680 | 30.797 | 1.00 | 25.30 | A | O |
| ATOM | 1298 | N | PRO | A | 180 | 75.586 | 20.866 | 32.358 | 1.00 | 25.86 | A | N |
| ATOM | 1299 | CD | PRO | A | 180 | 74.732 | 21.032 | 33.549 | 1.00 | 26.07 | A | C |
| ATOM | 1300 | CA | PRO | A | 180 | 75.424 | 21.988 | 31.426 | 1.00 | 26.24 | A | C |
| ATOM | 1301 | CB | PRO | A | 180 | 74.742 | 23.043 | 32.288 | 1.00 | 25.91 | A | C |
| ATOM | 1302 | CG | PRO | A | 180 | 73.829 | 22.189 | 33.141 | 1.00 | 27.50 | A | C |
| ATOM | 1303 | C | PRO | A | 180 | 76.762 | 22.460 | 30.843 | 1.00 | 26.29 | A | C |
| ATOM | 1304 | O | PRO | A | 180 | 76.845 | 22.823 | 29.668 | 1.00 | 26.65 | A | O |
| ATOM | 1305 | N | ALA | A | 181 | 77.808 | 22.447 | 31.661 | 1.00 | 26.19 | A | N |
| ATOM | 1306 | CA | ALA | A | 181 | 79.123 | 22.869 | 31.197 | 1.00 | 26.48 | A | C |
| ATOM | 1307 | CB | ALA | A | 181 | 80.139 | 22.743 | 32.307 | 1.00 | 26.08 | A | C |
| ATOM | 1308 | C | ALA | A | 181 | 79.549 | 22.024 | 30.004 | 1.00 | 26.39 | A | C |
| ATOM | 1309 | O | ALA | A | 181 | 79.820 | 22.553 | 28.925 | 1.00 | 27.82 | A | O |
| ATOM | 1310 | N | GLN | A | 182 | 79.602 | 20.712 | 30.199 | 1.00 | 25.31 | A | N |
| ATOM | 1311 | CA | GLN | A | 182 | 79.996 | 19.801 | 29.126 | 1.00 | 26.01 | A | C |
| ATOM | 1312 | CB | GLN | A | 182 | 79.845 | 18.353 | 29.595 | 1.00 | 26.47 | A | C |
| ATOM | 1313 | CG | GLN | A | 182 | 80.671 | 18.017 | 30.836 | 1.00 | 27.70 | A | C |
| ATOM | 1314 | CD | GLN | A | 182 | 80.371 | 16.624 | 31.355 | 1.00 | 29.50 | A | C |
| ATOM | 1315 | OE1 | GLN | A | 182 | 79.218 | 16.303 | 31.654 | 1.00 | 28.37 | A | O |
| ATOM | 1316 | NE2 | GLN | A | 182 | 81.408 | 15.783 | 31.461 | 1.00 | 28.58 | A | N |
| ATOM | 1317 | C | GLN | A | 182 | 79.180 | 20.021 | 27.845 | 1.00 | 25.57 | A | C |

FIGURE 9 - 20

SUBSTITUTE SHEET (RULE 26)

30/76

| | | | | | | | | | | |
|------|------|-----|-----------|--------|--------|--------|------|-------|---|---|
| ATOM | 1318 | O | GLN A 182 | 79.738 | 20.034 | 26.749 | 1.00 | 26.05 | A | O |
| ATOM | 1319 | N | ALA A 183 | 77.866 | 20.189 | 27.983 | 1.00 | 24.98 | A | N |
| ATOM | 1320 | CA | ALA A 183 | 76.999 | 20.412 | 26.829 | 1.00 | 25.67 | A | C |
| ATOM | 1321 | CB | ALA A 183 | 75.550 | 20.544 | 27.269 | 1.00 | 25.58 | A | C |
| ATOM | 1322 | C | ALA A 183 | 77.439 | 21.675 | 26.110 | 1.00 | 26.88 | A | C |
| ATOM | 1323 | O | ALA A 183 | 77.634 | 21.674 | 24.889 | 1.00 | 28.07 | A | O |
| ATOM | 1324 | N | GLN A 184 | 77.592 | 22.753 | 26.871 | 1.00 | 27.50 | A | N |
| ATOM | 1325 | CA | GLN A 184 | 78.029 | 24.032 | 26.312 | 1.00 | 27.76 | A | C |
| ATOM | 1326 | CB | GLN A 184 | 78.218 | 25.063 | 27.436 | 1.00 | 26.62 | A | C |
| ATOM | 1327 | CG | GLN A 184 | 79.011 | 26.321 | 27.072 | 1.00 | 25.37 | A | C |
| ATOM | 1328 | CD | GLN A 184 | 78.482 | 27.054 | 25.845 | 1.00 | 26.50 | A | C |
| ATOM | 1329 | OE1 | GLN A 184 | 77.271 | 27.187 | 25.650 | 1.00 | 26.90 | A | O |
| ATOM | 1330 | NE2 | GLN A 184 | 79.394 | 27.548 | 25.021 | 1.00 | 25.19 | A | N |
| ATOM | 1331 | C | GLN A 184 | 79.340 | 23.821 | 25.568 | 1.00 | 28.54 | A | C |
| ATOM | 1332 | O | GLN A 184 | 79.524 | 24.327 | 24.463 | 1.00 | 30.57 | A | O |
| ATOM | 1333 | N | GLU A 185 | 80.242 | 23.054 | 26.172 | 1.00 | 28.37 | A | N |
| ATOM | 1334 | CA | GLU A 185 | 81.531 | 22.788 | 25.563 | 1.00 | 27.94 | A | C |
| ATOM | 1335 | CB | GLU A 185 | 82.330 | 21.817 | 26.435 | 1.00 | 29.22 | A | C |
| ATOM | 1336 | CG | GLU A 185 | 83.761 | 21.608 | 25.958 | 1.00 | 32.97 | A | C |
| ATOM | 1337 | CD | GLU A 185 | 84.552 | 20.649 | 26.839 | 1.00 | 34.93 | A | C |
| ATOM | 1338 | OE1 | GLU A 185 | 85.771 | 20.497 | 26.604 | 1.00 | 36.75 | A | O |
| ATOM | 1339 | OE2 | GLU A 185 | 83.960 | 20.042 | 27.759 | 1.00 | 37.73 | A | O |
| ATOM | 1340 | C | GLU A 185 | 81.407 | 22.234 | 24.141 | 1.00 | 26.91 | A | C |
| ATOM | 1341 | O | GLU A 185 | 82.082 | 22.713 | 23.235 | 1.00 | 27.03 | A | O |
| ATOM | 1342 | N | VAL A 186 | 80.552 | 21.229 | 23.941 | 1.00 | 26.38 | A | N |
| ATOM | 1343 | CA | VAL A 186 | 80.379 | 20.644 | 22.611 | 1.00 | 23.65 | A | C |
| ATOM | 1344 | CB | VAL A 186 | 79.619 | 19.300 | 22.666 | 1.00 | 23.28 | A | C |
| ATOM | 1345 | CG1 | VAL A 186 | 79.531 | 18.707 | 21.281 | 1.00 | 23.35 | A | C |
| ATOM | 1346 | CG2 | VAL A 186 | 80.325 | 18.331 | 23.588 | 1.00 | 22.62 | A | C |
| ATOM | 1347 | C | VAL A 186 | 79.639 | 21.607 | 21.677 | 1.00 | 23.04 | A | C |
| ATOM | 1348 | O | VAL A 186 | 80.032 | 21.773 | 20.524 | 1.00 | 22.61 | A | O |
| ATOM | 1349 | N | HIS A 187 | 78.574 | 22.241 | 22.166 | 1.00 | 22.81 | A | N |
| ATOM | 1350 | CA | HIS A 187 | 77.832 | 23.192 | 21.337 | 1.00 | 22.39 | A | C |
| ATOM | 1351 | CB | HIS A 187 | 76.692 | 23.839 | 22.120 | 1.00 | 21.45 | A | C |
| ATOM | 1352 | CG | HIS A 187 | 75.482 | 22.968 | 22.260 | 1.00 | 22.41 | A | C |
| ATOM | 1353 | CD2 | HIS A 187 | 75.327 | 21.734 | 22.794 | 1.00 | 22.05 | A | C |
| ATOM | 1354 | ND1 | HIS A 187 | 74.228 | 23.367 | 21.850 | 1.00 | 22.74 | A | N |
| ATOM | 1355 | CE1 | HIS A 187 | 73.352 | 22.418 | 22.127 | 1.00 | 21.44 | A | C |
| ATOM | 1356 | NE2 | HIS A 187 | 73.993 | 21.416 | 22.700 | 1.00 | 23.45 | A | N |
| ATOM | 1357 | C | HIS A 187 | 78.772 | 24.275 | 20.824 | 1.00 | 23.56 | A | C |
| ATOM | 1358 | O | HIS A 187 | 78.729 | 24.635 | 19.653 | 1.00 | 23.80 | A | O |
| ATOM | 1359 | N | ALA A 188 | 79.635 | 24.781 | 21.699 | 1.00 | 24.61 | A | N |
| ATOM | 1360 | CA | ALA A 188 | 80.583 | 25.809 | 21.296 | 1.00 | 26.21 | A | C |
| ATOM | 1361 | CB | ALA A 188 | 81.414 | 26.264 | 22.498 | 1.00 | 26.29 | A | C |
| ATOM | 1362 | C | ALA A 188 | 81.492 | 25.248 | 20.212 | 1.00 | 26.14 | A | C |
| ATOM | 1363 | O | ALA A 188 | 81.704 | 25.871 | 19.172 | 1.00 | 27.35 | A | O |
| ATOM | 1364 | N | ALA A 189 | 82.020 | 24.056 | 20.452 | 1.00 | 26.39 | A | N |
| ATOM | 1365 | CA | ALA A 189 | 82.911 | 23.431 | 19.483 | 1.00 | 26.36 | A | C |
| ATOM | 1366 | CB | ALA A 189 | 83.457 | 22.134 | 20.040 | 1.00 | 27.40 | A | C |
| ATOM | 1367 | C | ALA A 189 | 82.199 | 23.175 | 18.161 | 1.00 | 26.38 | A | C |
| ATOM | 1368 | O | ALA A 189 | 82.817 | 23.250 | 17.094 | 1.00 | 27.20 | A | O |
| ATOM | 1369 | N | ILE A 190 | 80.903 | 22.871 | 18.220 | 1.00 | 25.44 | A | N |
| ATOM | 1370 | CA | ILE A 190 | 80.154 | 22.619 | 16.994 | 1.00 | 24.43 | A | C |
| ATOM | 1371 | CB | ILE A 190 | 78.735 | 22.083 | 17.279 | 1.00 | 23.66 | A | C |
| ATOM | 1372 | CG2 | ILE A 190 | 77.883 | 22.149 | 16.008 | 1.00 | 22.09 | A | C |
| ATOM | 1373 | CG1 | ILE A 190 | 78.827 | 20.640 | 17.771 | 1.00 | 21.03 | A | C |
| ATOM | 1374 | CD1 | ILE A 190 | 77.512 | 20.039 | 18.135 | 1.00 | 21.86 | A | C |
| ATOM | 1375 | C | ILE A 190 | 80.040 | 23.887 | 16.163 | 1.00 | 24.98 | A | C |
| ATOM | 1376 | O | ILE A 190 | 80.234 | 23.856 | 14.951 | 1.00 | 24.88 | A | O |
| ATOM | 1377 | N | ARG A 191 | 79.736 | 25.004 | 16.816 | 1.00 | 25.11 | A | N |
| ATOM | 1378 | CA | ARG A 191 | 79.601 | 26.272 | 16.108 | 1.00 | 25.21 | A | C |
| ATOM | 1379 | CB | ARG A 191 | 79.106 | 27.365 | 17.067 | 1.00 | 23.65 | A | C |
| ATOM | 1380 | CG | ARG A 191 | 78.394 | 28.517 | 16.374 | 1.00 | 21.44 | A | C |
| ATOM | 1381 | CD | ARG A 191 | 77.875 | 29.516 | 17.386 | 1.00 | 21.68 | A | C |
| ATOM | 1382 | NE | ARG A 191 | 76.857 | 28.947 | 18.262 | 1.00 | 21.18 | A | N |
| ATOM | 1383 | CZ | ARG A 191 | 75.583 | 28.770 | 17.916 | 1.00 | 22.25 | A | C |

FIGURE 9 - 21

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 1384 | NH1 | ARG | A | 191 | 75.157 | 29.119 | 16.705 | 1.00 | 18.54 | A | N |
| ATOM | 1385 | NH2 | ARG | A | 191 | 74.729 | 28.251 | 18.785 | 1.00 | 21.06 | A | N |
| ATOM | 1386 | C | ARG | A | 191 | 80.945 | 26.670 | 15.487 | 1.00 | 25.48 | A | C |
| ATOM | 1387 | O | ARG | A | 191 | 80.998 | 27.181 | 14.365 | 1.00 | 24.40 | A | O |
| ATOM | 1388 | N | ALA | A | 192 | 82.025 | 26.415 | 16.216 | 1.00 | 24.91 | A | N |
| ATOM | 1389 | CA | ALA | A | 192 | 83.360 | 26.738 | 15.735 | 1.00 | 26.61 | A | C |
| ATOM | 1390 | CB | ALA | A | 192 | 84.392 | 26.399 | 16.797 | 1.00 | 27.30 | A | C |
| ATOM | 1391 | C | ALA | A | 192 | 83.660 | 25.984 | 14.446 | 1.00 | 26.91 | A | C |
| ATOM | 1392 | O | ALA | A | 192 | 84.286 | 26.527 | 13.538 | 1.00 | 27.38 | A | O |
| ATOM | 1393 | N | GLN | A | 193 | 83.222 | 24.731 | 14.370 | 1.00 | 27.46 | A | N |
| ATOM | 1394 | CA | GLN | A | 193 | 83.433 | 23.933 | 13.164 | 1.00 | 27.39 | A | C |
| ATOM | 1395 | CB | GLN | A | 193 | 82.951 | 22.488 | 13.366 | 1.00 | 28.84 | A | C |
| ATOM | 1396 | CG | GLN | A | 193 | 83.725 | 21.743 | 14.451 | 1.00 | 32.03 | A | C |
| ATOM | 1397 | CD | GLN | A | 193 | 83.852 | 20.257 | 14.179 | 1.00 | 33.39 | A | C |
| ATOM | 1398 | OE1 | GLN | A | 193 | 84.506 | 19.528 | 14.927 | 1.00 | 34.59 | A | O |
| ATOM | 1399 | NE2 | GLN | A | 193 | 83.228 | 19.799 | 13.105 | 1.00 | 36.67 | A | N |
| ATOM | 1400 | C | GLN | A | 193 | 82.649 | 24.590 | 12.038 | 1.00 | 27.19 | A | C |
| ATOM | 1401 | O | GLN | A | 193 | 83.190 | 24.858 | 10.964 | 1.00 | 27.42 | A | O |
| ATOM | 1402 | N | LEU | A | 194 | 81.373 | 24.856 | 12.297 | 1.00 | 25.43 | A | N |
| ATOM | 1403 | CA | LEU | A | 194 | 80.526 | 25.504 | 11.313 | 1.00 | 24.72 | A | C |
| ATOM | 1404 | CB | LEU | A | 194 | 79.103 | 25.696 | 11.857 | 1.00 | 23.95 | A | C |
| ATOM | 1405 | CG | LEU | A | 194 | 78.202 | 24.457 | 12.004 | 1.00 | 25.27 | A | C |
| ATOM | 1406 | CD1 | LEU | A | 194 | 76.838 | 24.888 | 12.504 | 1.00 | 24.52 | A | C |
| ATOM | 1407 | CD2 | LEU | A | 194 | 78.050 | 23.747 | 10.660 | 1.00 | 25.88 | A | C |
| ATOM | 1408 | C | LEU | A | 194 | 81.131 | 26.854 | 10.951 | 1.00 | 23.26 | A | C |
| ATOM | 1409 | O | LEU | A | 194 | 81.105 | 27.256 | 9.794 | 1.00 | 24.92 | A | O |
| ATOM | 1410 | N | ALA | A | 195 | 81.693 | 27.537 | 11.943 | 1.00 | 24.08 | A | N |
| ATOM | 1411 | CA | ALA | A | 195 | 82.304 | 28.854 | 11.746 | 1.00 | 24.66 | A | C |
| ATOM | 1412 | CB | ALA | A | 195 | 82.653 | 29.476 | 13.110 | 1.00 | 25.98 | A | C |
| ATOM | 1413 | C | ALA | A | 195 | 83.537 | 28.869 | 10.831 | 1.00 | 24.98 | A | C |
| ATOM | 1414 | O | ALA | A | 195 | 83.760 | 29.841 | 10.111 | 1.00 | 23.53 | A | O |
| ATOM | 1415 | N | ALA | A | 196 | 84.336 | 27.805 | 10.860 | 1.00 | 24.75 | A | N |
| ATOM | 1416 | CA | ALA | A | 196 | 85.524 | 27.735 | 10.002 | 1.00 | 26.30 | A | C |
| ATOM | 1417 | CB | ALA | A | 196 | 86.381 | 26.518 | 10.375 | 1.00 | 25.46 | A | C |
| ATOM | 1418 | C | ALA | A | 196 | 85.123 | 27.658 | 8.519 | 1.00 | 26.39 | A | C |
| ATOM | 1419 | O | ALA | A | 196 | 85.923 | 27.966 | 7.634 | 1.00 | 27.18 | A | O |
| ATOM | 1420 | N | GLU | A | 197 | 83.890 | 27.242 | 8.246 | 1.00 | 26.05 | A | N |
| ATOM | 1421 | CA | GLU | A | 197 | 83.431 | 27.149 | 6.863 | 1.00 | 25.61 | A | C |
| ATOM | 1422 | CB | GLU | A | 197 | 82.464 | 25.968 | 6.701 | 1.00 | 24.27 | A | C |
| ATOM | 1423 | CG | GLU | A | 197 | 83.048 | 24.616 | 7.158 | 1.00 | 25.07 | A | C |
| ATOM | 1424 | CD | GLU | A | 197 | 84.274 | 24.156 | 6.353 | 1.00 | 23.08 | A | C |
| ATOM | 1425 | OE1 | GLU | A | 197 | 85.253 | 23.697 | 6.976 | 1.00 | 22.55 | A | O |
| ATOM | 1426 | OE2 | GLU | A | 197 | 84.260 | 24.231 | 5.107 | 1.00 | 24.35 | A | O |
| ATOM | 1427 | C | GLU | A | 197 | 82.766 | 28.468 | 6.447 | 1.00 | 26.00 | A | C |
| ATOM | 1428 | O | GLU | A | 197 | 82.937 | 28.924 | 5.320 | 1.00 | 27.03 | A | O |
| ATOM | 1429 | N | ASN | A | 198 | 82.015 | 29.074 | 7.362 | 1.00 | 25.11 | A | N |
| ATOM | 1430 | CA | ASN | A | 198 | 81.354 | 30.358 | 7.114 | 1.00 | 25.09 | A | C |
| ATOM | 1431 | CB | ASN | A | 198 | 80.123 | 30.199 | 6.212 | 1.00 | 23.56 | A | C |
| ATOM | 1432 | CG | ASN | A | 198 | 79.476 | 31.550 | 5.849 | 1.00 | 25.67 | A | C |
| ATOM | 1433 | OD1 | ASN | A | 198 | 79.012 | 32.292 | 6.725 | 1.00 | 25.82 | A | O |
| ATOM | 1434 | ND2 | ASN | A | 198 | 79.448 | 31.865 | 4.553 | 1.00 | 22.03 | A | N |
| ATOM | 1435 | C | ASN | A | 198 | 80.930 | 30.925 | 8.464 | 1.00 | 25.86 | A | C |
| ATOM | 1436 | O | ASN | A | 198 | 79.975 | 30.448 | 9.073 | 1.00 | 24.71 | A | O |
| ATOM | 1437 | N | ALA | A | 199 | 81.650 | 31.942 | 8.926 | 1.00 | 27.00 | A | N |
| ATOM | 1438 | CA | ALA | A | 199 | 81.354 | 32.551 | 10.217 | 1.00 | 29.32 | A | C |
| ATOM | 1439 | CB | ALA | A | 199 | 82.371 | 33.648 | 10.518 | 1.00 | 29.39 | A | C |
| ATOM | 1440 | C | ALA | A | 199 | 79.929 | 33.104 | 10.346 | 1.00 | 29.29 | A | C |
| ATOM | 1441 | O | ALA | A | 199 | 79.259 | 32.855 | 11.348 | 1.00 | 30.16 | A | O |
| ATOM | 1442 | N | GLU | A | 200 | 79.459 | 33.842 | 9.346 | 1.00 | 29.71 | A | N |
| ATOM | 1443 | CA | GLU | A | 200 | 78.116 | 34.406 | 9.422 | 1.00 | 31.28 | A | C |
| ATOM | 1444 | CB | GLU | A | 200 | 77.771 | 35.234 | 8.173 | 1.00 | 34.61 | A | C |
| ATOM | 1445 | CG | GLU | A | 200 | 76.444 | 35.985 | 8.341 | 1.00 | 39.99 | A | C |
| ATOM | 1446 | CD | GLU | A | 200 | 76.172 | 37.031 | 7.263 | 1.00 | 45.82 | A | C |
| ATOM | 1447 | OE1 | GLU | A | 200 | 77.074 | 37.864 | 6.983 | 1.00 | 46.88 | A | O |
| ATOM | 1448 | OE2 | GLU | A | 200 | 75.041 | 37.038 | 6.704 | 1.00 | 47.99 | A | O |
| ATOM | 1449 | C | GLU | A | 200 | 77.059 | 33.325 | 9.619 | 1.00 | 30.52 | A | C |

FIGURE 9 - 22

SUBSTITUTE SHEET (RULE 26)

32/76

| | | | | | | | | | | |
|------|------|-----|-----------|--------|--------|--------|------|-------|---|---|
| ATOM | 1450 | O | GLU A 200 | 76.230 | 33.419 | 10.523 | 1.00 | 31.49 | A | O |
| ATOM | 1451 | N | VAL A 201 | 77.096 | 32.292 | 8.786 | 1.00 | 28.73 | A | N |
| ATOM | 1452 | CA | VAL A 201 | 76.133 | 31.206 | 8.894 | 1.00 | 27.29 | A | C |
| ATOM | 1453 | CB | VAL A 201 | 76.413 | 30.120 | 7.834 | 1.00 | 28.50 | A | C |
| ATOM | 1454 | CG1 | VAL A 201 | 75.519 | 28.896 | 8.083 | 1.00 | 29.52 | A | C |
| ATOM | 1455 | CG2 | VAL A 201 | 76.163 | 30.695 | 6.435 | 1.00 | 25.40 | A | C |
| ATOM | 1456 | C | VAL A 201 | 76.164 | 30.569 | 10.281 | 1.00 | 26.11 | A | C |
| ATOM | 1457 | O | VAL A 201 | 75.123 | 30.270 | 10.862 | 1.00 | 26.64 | A | O |
| ATOM | 1458 | N | ALA A 202 | 77.365 | 30.375 | 10.808 | 1.00 | 25.34 | A | N |
| ATOM | 1459 | CA | ALA A 202 | 77.543 | 29.763 | 12.117 | 1.00 | 25.05 | A | C |
| ATOM | 1460 | CB | ALA A 202 | 79.033 | 29.564 | 12.391 | 1.00 | 22.78 | A | C |
| ATOM | 1461 | C | ALA A 202 | 76.897 | 30.550 | 13.261 | 1.00 | 24.66 | A | C |
| ATOM | 1462 | O | ALA A 202 | 76.404 | 29.964 | 14.220 | 1.00 | 24.39 | A | O |
| ATOM | 1463 | N | LYS A 203 | 76.901 | 31.872 | 13.162 | 1.00 | 24.28 | A | N |
| ATOM | 1464 | CA | LYS A 203 | 76.320 | 32.699 | 14.210 | 1.00 | 25.27 | A | C |
| ATOM | 1465 | CB | LYS A 203 | 76.766 | 34.152 | 14.049 | 1.00 | 27.11 | A | C |
| ATOM | 1466 | CG | LYS A 203 | 78.258 | 34.361 | 14.204 | 1.00 | 31.07 | A | C |
| ATOM | 1467 | CD | LYS A 203 | 78.664 | 35.800 | 13.848 | 1.00 | 35.87 | A | C |
| ATOM | 1468 | CE | LYS A 203 | 80.189 | 35.972 | 13.909 | 1.00 | 37.34 | A | C |
| ATOM | 1469 | NZ | LYS A 203 | 80.627 | 37.296 | 13.360 | 1.00 | 40.28 | A | N |
| ATOM | 1470 | C | LYS A 203 | 74.804 | 32.637 | 14.205 | 1.00 | 24.60 | A | C |
| ATOM | 1471 | O | LYS A 203 | 74.166 | 32.913 | 15.220 | 1.00 | 23.69 | A | O |
| ATOM | 1472 | N | GLY A 204 | 74.234 | 32.261 | 13.065 | 1.00 | 24.21 | A | N |
| ATOM | 1473 | CA | GLY A 204 | 72.792 | 32.189 | 12.957 | 1.00 | 24.41 | A | C |
| ATOM | 1474 | C | GLY A 204 | 72.149 | 30.832 | 13.173 | 1.00 | 23.93 | A | C |
| ATOM | 1475 | O | GLY A 204 | 70.958 | 30.758 | 13.440 | 1.00 | 24.78 | A | O |
| ATOM | 1476 | N | VAL A 205 | 72.920 | 29.758 | 13.077 | 1.00 | 23.36 | A | N |
| ATOM | 1477 | CA | VAL A 205 | 72.355 | 28.423 | 13.239 | 1.00 | 23.48 | A | C |
| ATOM | 1478 | CB | VAL A 205 | 73.350 | 27.337 | 12.752 | 1.00 | 24.84 | A | C |
| ATOM | 1479 | CG1 | VAL A 205 | 74.635 | 27.397 | 13.572 | 1.00 | 23.10 | A | C |
| ATOM | 1480 | CG2 | VAL A 205 | 72.703 | 25.958 | 12.873 | 1.00 | 21.54 | A | C |
| ATOM | 1481 | C | VAL A 205 | 71.920 | 28.050 | 14.658 | 1.00 | 22.77 | A | C |
| ATOM | 1482 | O | VAL A 205 | 72.721 | 28.082 | 15.580 | 1.00 | 23.72 | A | O |
| ATOM | 1483 | N | ARG A 206 | 70.655 | 27.682 | 14.837 | 1.00 | 22.07 | A | N |
| ATOM | 1484 | CA | ARG A 206 | 70.206 | 27.280 | 16.166 | 1.00 | 22.57 | A | C |
| ATOM | 1485 | CB | ARG A 206 | 68.675 | 27.248 | 16.255 | 1.00 | 23.65 | A | C |
| ATOM | 1486 | CG | ARG A 206 | 68.029 | 28.621 | 16.406 | 1.00 | 25.28 | A | C |
| ATOM | 1487 | CD | ARG A 206 | 67.329 | 29.070 | 15.137 | 1.00 | 27.42 | A | C |
| ATOM | 1488 | NE | ARG A 206 | 65.894 | 28.795 | 15.178 | 1.00 | 30.23 | A | N |
| ATOM | 1489 | CZ | ARG A 206 | 65.183 | 28.373 | 14.132 | 1.00 | 31.92 | A | C |
| ATOM | 1490 | NH1 | ARG A 206 | 65.772 | 28.172 | 12.950 | 1.00 | 30.96 | A | N |
| ATOM | 1491 | NH2 | ARG A 206 | 63.881 | 28.146 | 14.268 | 1.00 | 31.21 | A | N |
| ATOM | 1492 | C | ARG A 206 | 70.773 | 25.900 | 16.499 | 1.00 | 21.98 | A | C |
| ATOM | 1493 | O | ARG A 206 | 70.637 | 24.963 | 15.715 | 1.00 | 20.78 | A | O |
| ATOM | 1494 | N | LEU A 207 | 71.443 | 25.792 | 17.644 | 1.00 | 21.95 | A | N |
| ATOM | 1495 | CA | LEU A 207 | 72.007 | 24.514 | 18.093 | 1.00 | 22.08 | A | C |
| ATOM | 1496 | CB | LEU A 207 | 73.489 | 24.660 | 18.461 | 1.00 | 19.71 | A | C |
| ATOM | 1497 | CG | LEU A 207 | 74.490 | 25.127 | 17.394 | 1.00 | 20.62 | A | C |
| ATOM | 1498 | CD1 | LEU A 207 | 75.895 | 25.142 | 18.004 | 1.00 | 20.74 | A | C |
| ATOM | 1499 | CD2 | LEU A 207 | 74.443 | 24.201 | 16.177 | 1.00 | 17.86 | A | C |
| ATOM | 1500 | C | LEU A 207 | 71.217 | 24.080 | 19.325 | 1.00 | 22.22 | A | C |
| ATOM | 1501 | O | LEU A 207 | 71.295 | 24.725 | 20.371 | 1.00 | 22.75 | A | O |
| ATOM | 1502 | N | LEU A 208 | 70.462 | 22.990 | 19.204 | 1.00 | 21.40 | A | N |
| ATOM | 1503 | CA | LEU A 208 | 69.643 | 22.509 | 20.313 | 1.00 | 21.08 | A | C |
| ATOM | 1504 | CB | LEU A 208 | 68.300 | 21.997 | 19.792 | 1.00 | 21.22 | A | C |
| ATOM | 1505 | CG | LEU A 208 | 67.640 | 22.665 | 18.583 | 1.00 | 21.15 | A | C |
| ATOM | 1506 | CD1 | LEU A 208 | 66.307 | 21.983 | 18.353 | 1.00 | 20.46 | A | C |
| ATOM | 1507 | CD2 | LEU A 208 | 67.438 | 24.158 | 18.806 | 1.00 | 22.79 | A | C |
| ATOM | 1508 | C | LEU A 208 | 70.300 | 21.387 | 21.110 | 1.00 | 22.11 | A | C |
| ATOM | 1509 | O | LEU A 208 | 71.001 | 20.547 | 20.547 | 1.00 | 21.35 | A | O |
| ATOM | 1510 | N | TYR A 209 | 70.065 | 21.377 | 22.421 | 1.00 | 19.68 | A | N |
| ATOM | 1511 | CA | TYR A 209 | 70.603 | 20.328 | 23.270 | 1.00 | 19.72 | A | C |
| ATOM | 1512 | CB | TYR A 209 | 70.756 | 20.810 | 24.711 | 1.00 | 19.55 | A | C |
| ATOM | 1513 | CG | TYR A 209 | 71.267 | 19.728 | 25.628 | 1.00 | 19.86 | A | C |
| ATOM | 1514 | CD1 | TYR A 209 | 72.533 | 19.168 | 25.435 | 1.00 | 19.09 | A | C |
| ATOM | 1515 | CE1 | TYR A 209 | 72.997 | 18.142 | 26.247 | 1.00 | 20.71 | A | C |

FIGURE 9 - 23

SUBSTITUTE SHEET (RULE 26)

33/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 1516 | CD2 | TYR | A | 209 | 70.475 | 19.232 | 26.662 | 1.00 | 18.33 | A | C |
| ATOM | 1517 | CE2 | TYR | A | 209 | 70.927 | 18.201 | 27.484 | 1.00 | 20.46 | A | C |
| ATOM | 1518 | CZ | TYR | A | 209 | 72.185 | 17.659 | 27.270 | 1.00 | 22.30 | A | C |
| ATOM | 1519 | OH | TYR | A | 209 | 72.619 | 16.618 | 28.049 | 1.00 | 25.22 | A | C |
| ATOM | 1520 | C | TYR | A | 209 | 69.625 | 19.153 | 23.227 | 1.00 | 20.01 | A | O |
| ATOM | 1521 | O | TYR | A | 209 | 68.426 | 19.329 | 23.465 | 1.00 | 18.91 | A | O |
| ATOM | 1522 | N | GLY | A | 210 | 70.141 | 17.961 | 22.940 | 1.00 | 18.48 | A | N |
| ATOM | 1523 | CA | GLY | A | 210 | 69.287 | 16.790 | 22.842 | 1.00 | 22.41 | A | C |
| ATOM | 1524 | C | GLY | A | 210 | 69.450 | 15.741 | 23.918 | 1.00 | 22.71 | A | C |
| ATOM | 1525 | O | GLY | A | 210 | 69.101 | 14.585 | 23.709 | 1.00 | 22.58 | A | O |
| ATOM | 1526 | N | GLY | A | 211 | 70.001 | 16.138 | 25.062 | 1.00 | 24.79 | A | N |
| ATOM | 1527 | CA | GLY | A | 211 | 70.177 | 15.209 | 26.167 | 1.00 | 26.36 | A | C |
| ATOM | 1528 | C | GLY | A | 211 | 68.929 | 15.236 | 27.029 | 1.00 | 27.13 | A | C |
| ATOM | 1529 | O | GLY | A | 211 | 67.875 | 15.663 | 26.556 | 1.00 | 26.25 | A | O |
| ATOM | 1530 | N | SER | A | 212 | 69.030 | 14.777 | 28.278 | 1.00 | 27.69 | A | N |
| ATOM | 1531 | CA | SER | A | 212 | 67.883 | 14.774 | 29.186 | 1.00 | 29.84 | A | C |
| ATOM | 1532 | CB | SER | A | 212 | 68.233 | 14.098 | 30.511 | 1.00 | 29.31 | A | C |
| ATOM | 1533 | OG | SER | A | 212 | 68.637 | 12.756 | 30.327 | 1.00 | 33.66 | A | O |
| ATOM | 1534 | C | SER | A | 212 | 67.450 | 16.206 | 29.484 | 1.00 | 30.00 | A | C |
| ATOM | 1535 | O | SER | A | 212 | 68.259 | 17.018 | 29.923 | 1.00 | 28.67 | A | O |
| ATOM | 1536 | N | VAL | A | 213 | 66.182 | 16.515 | 29.239 | 1.00 | 30.24 | A | N |
| ATOM | 1537 | CA | VAL | A | 213 | 65.671 | 17.853 | 29.513 | 1.00 | 30.55 | A | C |
| ATOM | 1538 | CB | VAL | A | 213 | 65.441 | 18.671 | 28.218 | 1.00 | 31.41 | A | C |
| ATOM | 1539 | CG1 | VAL | A | 213 | 64.905 | 20.064 | 28.574 | 1.00 | 31.49 | A | C |
| ATOM | 1540 | CG2 | VAL | A | 213 | 66.741 | 18.795 | 27.425 | 1.00 | 31.44 | A | C |
| ATOM | 1541 | C | VAL | A | 213 | 64.336 | 17.767 | 30.224 | 1.00 | 30.96 | A | C |
| ATOM | 1542 | O | VAL | A | 213 | 63.386 | 17.201 | 29.687 | 1.00 | 32.12 | A | O |
| ATOM | 1543 | N | LYS | A | 214 | 64.264 | 18.321 | 31.432 | 1.00 | 30.80 | A | N |
| ATOM | 1544 | CA | LYS | A | 214 | 63.015 | 18.344 | 32.186 | 1.00 | 30.16 | A | C |
| ATOM | 1545 | CB | LYS | A | 214 | 63.066 | 17.379 | 33.375 | 1.00 | 31.65 | A | C |
| ATOM | 1546 | CG | LYS | A | 214 | 64.219 | 17.592 | 34.327 | 1.00 | 33.28 | A | C |
| ATOM | 1547 | CD | LYS | A | 214 | 64.217 | 16.527 | 35.417 | 1.00 | 35.88 | A | C |
| ATOM | 1548 | CE | LYS | A | 214 | 64.333 | 15.135 | 34.812 | 1.00 | 38.58 | A | C |
| ATOM | 1549 | NZ | LYS | A | 214 | 64.467 | 14.061 | 35.851 | 1.00 | 40.77 | A | N |
| ATOM | 1550 | C | LYS | A | 214 | 62.712 | 19.763 | 32.661 | 1.00 | 29.23 | A | C |
| ATOM | 1551 | O | LYS | A | 214 | 63.548 | 20.664 | 32.542 | 1.00 | 27.70 | A | O |
| ATOM | 1552 | N | ALA | A | 215 | 61.509 | 19.957 | 33.191 | 1.00 | 28.62 | A | N |
| ATOM | 1553 | CA | ALA | A | 215 | 61.071 | 21.269 | 33.658 | 1.00 | 28.86 | A | C |
| ATOM | 1554 | CB | ALA | A | 215 | 59.722 | 21.149 | 34.360 | 1.00 | 27.98 | A | C |
| ATOM | 1555 | C | ALA | A | 215 | 62.073 | 21.931 | 34.588 | 1.00 | 27.64 | A | C |
| ATOM | 1556 | O | ALA | A | 215 | 62.409 | 23.102 | 34.424 | 1.00 | 28.09 | A | O |
| ATOM | 1557 | N | ALA | A | 216 | 62.559 | 21.173 | 35.558 | 1.00 | 27.89 | A | N |
| ATOM | 1558 | CA | ALA | A | 216 | 63.497 | 21.702 | 36.542 | 1.00 | 28.54 | A | C |
| ATOM | 1559 | CB | ALA | A | 216 | 63.490 | 20.803 | 37.779 | 1.00 | 28.46 | A | C |
| ATOM | 1560 | C | ALA | A | 216 | 64.938 | 21.924 | 36.079 | 1.00 | 27.78 | A | C |
| ATOM | 1561 | O | ALA | A | 216 | 65.715 | 22.564 | 36.785 | 1.00 | 27.58 | A | O |
| ATOM | 1562 | N | SER | A | 217 | 65.297 | 21.418 | 34.901 | 1.00 | 28.01 | A | N |
| ATOM | 1563 | CA | SER | A | 217 | 66.671 | 21.561 | 34.409 | 1.00 | 27.38 | A | C |
| ATOM | 1564 | CB | SER | A | 217 | 67.261 | 20.182 | 34.108 | 1.00 | 26.95 | A | C |
| ATOM | 1565 | OG | SER | A | 217 | 66.542 | 19.561 | 33.054 | 1.00 | 24.95 | A | O |
| ATOM | 1566 | C | SER | A | 217 | 66.808 | 22.409 | 33.156 | 1.00 | 27.32 | A | C |
| ATOM | 1567 | O | SER | A | 217 | 67.902 | 22.876 | 32.835 | 1.00 | 26.32 | A | O |
| ATOM | 1568 | N | ALA | A | 218 | 65.694 | 22.603 | 32.456 | 1.00 | 26.76 | A | N |
| ATOM | 1569 | CA | ALA | A | 218 | 65.680 | 23.355 | 31.206 | 1.00 | 25.53 | A | C |
| ATOM | 1570 | CB | ALA | A | 218 | 64.255 | 23.415 | 30.659 | 1.00 | 25.02 | A | C |
| ATOM | 1571 | C | ALA | A | 218 | 66.274 | 24.760 | 31.258 | 1.00 | 25.94 | A | C |
| ATOM | 1572 | O | ALA | A | 218 | 67.049 | 25.136 | 30.381 | 1.00 | 26.29 | A | O |
| ATOM | 1573 | N | ALA | A | 219 | 65.922 | 25.529 | 32.279 | 1.00 | 24.83 | A | N |
| ATOM | 1574 | CA | ALA | A | 219 | 66.408 | 26.902 | 32.384 | 1.00 | 25.74 | A | C |
| ATOM | 1575 | CB | ALA | A | 219 | 65.843 | 27.552 | 33.638 | 1.00 | 25.62 | A | C |
| ATOM | 1576 | C | ALA | A | 219 | 67.928 | 27.057 | 32.354 | 1.00 | 25.11 | A | C |
| ATOM | 1577 | O | ALA | A | 219 | 68.453 | 27.873 | 31.597 | 1.00 | 24.31 | A | O |
| ATOM | 1578 | N | GLU | A | 220 | 68.619 | 26.278 | 33.185 | 1.00 | 26.90 | A | N |
| ATOM | 1579 | CA | GLU | A | 220 | 70.078 | 26.313 | 33.294 | 1.00 | 26.66 | A | C |
| ATOM | 1580 | CB | GLU | A | 220 | 70.524 | 25.369 | 34.426 | 1.00 | 29.38 | A | C |
| ATOM | 1581 | CG | GLU | A | 220 | 72.030 | 25.071 | 34.528 | 1.00 | 32.92 | A | C |

FIGURE 9- 24

SUBSTITUTE SHEET (RULE 26)

34/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 1582 | CD | GLU | A | 220 | 72.407 | 24.398 | 35.863 | 1.00 | 36.72 | A | C |
| ATOM | 1583 | OE1 | GLU | A | 220 | 71.632 | 23.545 | 36.357 | 1.00 | 39.37 | A | O |
| ATOM | 1584 | OE2 | GLU | A | 220 | 73.484 | 24.708 | 36.424 | 1.00 | 37.71 | A | O |
| ATOM | 1585 | C | GLU | A | 220 | 70.739 | 25.930 | 31.975 | 1.00 | 27.18 | A | C |
| ATOM | 1586 | O | GLU | A | 220 | 71.684 | 26.580 | 31.524 | 1.00 | 27.22 | A | O |
| ATOM | 1587 | N | LEU | A | 221 | 70.237 | 24.872 | 31.354 | 1.00 | 26.04 | A | N |
| ATOM | 1588 | CA | LEU | A | 221 | 70.781 | 24.421 | 30.087 | 1.00 | 25.54 | A | C |
| ATOM | 1589 | CB | LEU | A | 221 | 70.154 | 23.072 | 29.711 | 1.00 | 25.36 | A | C |
| ATOM | 1590 | CG | LEU | A | 221 | 70.814 | 21.831 | 30.323 | 1.00 | 24.33 | A | C |
| ATOM | 1591 | CD1 | LEU | A | 221 | 69.912 | 20.623 | 30.183 | 1.00 | 25.52 | A | C |
| ATOM | 1592 | CD2 | LEU | A | 221 | 72.152 | 21.585 | 29.635 | 1.00 | 23.47 | A | C |
| ATOM | 1593 | C | LEU | A | 221 | 70.561 | 25.454 | 28.971 | 1.00 | 24.88 | A | C |
| ATOM | 1594 | O | LEU | A | 221 | 71.508 | 25.868 | 28.303 | 1.00 | 22.01 | A | O |
| ATOM | 1595 | N | PHE | A | 222 | 69.311 | 25.874 | 28.784 | 1.00 | 25.62 | A | N |
| ATOM | 1596 | CA | PHE | A | 222 | 68.977 | 26.839 | 27.743 | 1.00 | 25.16 | A | C |
| ATOM | 1597 | CB | PHE | A | 222 | 67.467 | 27.086 | 27.702 | 1.00 | 23.47 | A | C |
| ATOM | 1598 | CG | PHE | A | 222 | 66.649 | 25.868 | 27.342 | 1.00 | 22.39 | A | C |
| ATOM | 1599 | CD1 | PHE | A | 222 | 67.229 | 24.781 | 26.690 | 1.00 | 20.09 | A | C |
| ATOM | 1600 | CD2 | PHE | A | 222 | 65.278 | 25.832 | 27.616 | 1.00 | 21.71 | A | C |
| ATOM | 1601 | CE1 | PHE | A | 222 | 66.459 | 23.676 | 26.314 | 1.00 | 20.60 | A | C |
| ATOM | 1602 | CE2 | PHE | A | 222 | 64.496 | 24.732 | 27.245 | 1.00 | 21.98 | A | C |
| ATOM | 1603 | CZ | PHE | A | 222 | 65.089 | 23.650 | 26.591 | 1.00 | 20.28 | A | C |
| ATOM | 1604 | C | PHE | A | 222 | 69.699 | 28.159 | 27.965 | 1.00 | 26.92 | A | C |
| ATOM | 1605 | O | PHE | A | 222 | 69.901 | 28.940 | 27.026 | 1.00 | 26.76 | A | O |
| ATOM | 1606 | N | GLY | A | 223 | 70.093 | 28.396 | 29.210 | 1.00 | 26.78 | A | N |
| ATOM | 1607 | CA | GLY | A | 223 | 70.783 | 29.624 | 29.544 | 1.00 | 27.48 | A | C |
| ATOM | 1608 | C | GLY | A | 223 | 72.214 | 29.674 | 29.043 | 1.00 | 27.70 | A | C |
| ATOM | 1609 | O | GLY | A | 223 | 72.808 | 30.744 | 28.951 | 1.00 | 28.93 | A | O |
| ATOM | 1610 | N | MET | A | 224 | 72.795 | 28.528 | 28.724 | 1.00 | 27.35 | A | N |
| ATOM | 1611 | CA | MET | A | 224 | 74.163 | 28.551 | 28.241 | 1.00 | 26.38 | A | C |
| ATOM | 1612 | CB | MET | A | 224 | 74.683 | 27.125 | 28.087 | 1.00 | 26.05 | A | C |
| ATOM | 1613 | CG | MET | A | 224 | 75.971 | 26.877 | 28.870 | 1.00 | 28.39 | A | C |
| ATOM | 1614 | SD | MET | A | 224 | 75.806 | 27.008 | 30.649 | 1.00 | 24.91 | A | S |
| ATOM | 1615 | CE | MET | A | 224 | 77.257 | 26.125 | 31.131 | 1.00 | 27.81 | A | C |
| ATOM | 1616 | C | MET | A | 224 | 74.241 | 29.334 | 26.913 | 1.00 | 25.44 | A | C |
| ATOM | 1617 | O | MET | A | 224 | 73.301 | 29.337 | 26.117 | 1.00 | 24.05 | A | O |
| ATOM | 1618 | N | PRO | A | 225 | 75.366 | 30.024 | 26.673 | 1.00 | 26.30 | A | N |
| ATOM | 1619 | CD | PRO | A | 225 | 76.582 | 30.042 | 27.504 | 1.00 | 27.39 | A | C |
| ATOM | 1620 | CA | PRO | A | 225 | 75.564 | 30.817 | 25.456 | 1.00 | 26.61 | A | C |
| ATOM | 1621 | CB | PRO | A | 225 | 76.982 | 31.361 | 25.625 | 1.00 | 28.19 | A | C |
| ATOM | 1622 | CG | PRO | A | 225 | 77.652 | 30.322 | 26.484 | 1.00 | 28.50 | A | C |
| ATOM | 1623 | C | PRO | A | 225 | 75.365 | 30.129 | 24.099 | 1.00 | 25.97 | A | C |
| ATOM | 1624 | O | PRO | A | 225 | 74.741 | 30.694 | 23.209 | 1.00 | 25.42 | A | O |
| ATOM | 1625 | N | ASP | A | 226 | 75.876 | 28.918 | 23.932 | 1.00 | 25.30 | A | N |
| ATOM | 1626 | CA | ASP | A | 226 | 75.733 | 28.256 | 22.641 | 1.00 | 26.40 | A | C |
| ATOM | 1627 | CB | ASP | A | 226 | 77.083 | 27.657 | 22.232 | 1.00 | 25.52 | A | C |
| ATOM | 1628 | CG | ASP | A | 226 | 78.132 | 28.732 | 21.956 | 1.00 | 28.90 | A | C |
| ATOM | 1629 | OD1 | ASP | A | 226 | 78.042 | 29.393 | 20.897 | 1.00 | 29.10 | A | O |
| ATOM | 1630 | OD2 | ASP | A | 226 | 79.038 | 28.926 | 22.803 | 1.00 | 29.86 | A | O |
| ATOM | 1631 | C | ASP | A | 226 | 74.603 | 27.218 | 22.531 | 1.00 | 25.89 | A | C |
| ATOM | 1632 | O | ASP | A | 226 | 74.570 | 26.416 | 21.595 | 1.00 | 23.67 | A | O |
| ATOM | 1633 | N | ILE | A | 227 | 73.675 | 27.249 | 23.484 | 1.00 | 24.91 | A | N |
| ATOM | 1634 | CA | ILE | A | 227 | 72.537 | 26.336 | 23.466 | 1.00 | 26.34 | A | C |
| ATOM | 1635 | CB | ILE | A | 227 | 72.366 | 25.623 | 24.813 | 1.00 | 27.46 | A | C |
| ATOM | 1636 | CG2 | ILE | A | 227 | 71.213 | 24.623 | 24.723 | 1.00 | 26.87 | A | C |
| ATOM | 1637 | CG1 | ILE | A | 227 | 73.669 | 24.913 | 25.190 | 1.00 | 27.28 | A | C |
| ATOM | 1638 | CD1 | ILE | A | 227 | 73.574 | 24.071 | 26.453 | 1.00 | 27.14 | A | C |
| ATOM | 1639 | C | ILE | A | 227 | 71.287 | 27.154 | 23.154 | 1.00 | 25.41 | A | C |
| ATOM | 1640 | O | ILE | A | 227 | 70.873 | 27.999 | 23.939 | 1.00 | 27.14 | A | O |
| ATOM | 1641 | N | ASP | A | 228 | 70.685 | 26.893 | 22.003 | 1.00 | 25.95 | A | N |
| ATOM | 1642 | CA | ASP | A | 228 | 69.524 | 27.651 | 21.562 | 1.00 | 25.76 | A | C |
| ATOM | 1643 | CB | ASP | A | 228 | 69.687 | 27.963 | 20.076 | 1.00 | 24.22 | A | C |
| ATOM | 1644 | CG | ASP | A | 228 | 71.015 | 28.639 | 19.776 | 1.00 | 27.22 | A | C |
| ATOM | 1645 | OD1 | ASP | A | 228 | 71.770 | 28.127 | 18.914 | 1.00 | 27.25 | A | O |
| ATOM | 1646 | OD2 | ASP | A | 228 | 71.302 | 29.681 | 20.413 | 1.00 | 26.31 | A | O |
| ATOM | 1647 | C | ASP | A | 228 | 68.160 | 27.020 | 21.818 | 1.00 | 25.91 | A | C |

FIGURE 9 - 25

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | | |
|------|------|-----|-----------|--------|--------|--------|------|-------|---|---|
| ATOM | 1648 | O | ASP A 228 | 67.149 | 27.471 | 21.273 | 1.00 | 27.55 | A | O |
| ATOM | 1649 | N | GLY A 229 | 68.128 | 25.995 | 22.659 | 1.00 | 24.49 | A | N |
| ATOM | 1650 | CA | GLY A 229 | 66.876 | 25.335 | 22.962 | 1.00 | 24.44 | A | C |
| ATOM | 1651 | C | GLY A 229 | 67.110 | 23.852 | 23.149 | 1.00 | 24.31 | A | C |
| ATOM | 1652 | O | GLY A 229 | 68.244 | 23.419 | 23.375 | 1.00 | 24.14 | A | O |
| ATOM | 1653 | N | GLY A 230 | 66.048 | 23.063 | 23.063 | 1.00 | 22.84 | A | N |
| ATOM | 1654 | CA | GLY A 230 | 66.213 | 21.636 | 23.236 | 1.00 | 22.20 | A | C |
| ATOM | 1655 | C | GLY A 230 | 65.335 | 20.759 | 22.360 | 1.00 | 23.21 | A | C |
| ATOM | 1656 | O | GLY A 230 | 64.293 | 21.190 | 21.843 | 1.00 | 22.77 | A | O |
| ATOM | 1657 | N | LEU A 231 | 65.789 | 19.525 | 22.173 | 1.00 | 21.78 | A | N |
| ATOM | 1658 | CA | LEU A 231 | 65.058 | 18.533 | 21.407 | 1.00 | 22.79 | A | C |
| ATOM | 1659 | CB | LEU A 231 | 65.972 | 17.837 | 20.394 | 1.00 | 20.20 | A | C |
| ATOM | 1660 | CG | LEU A 231 | 65.335 | 16.753 | 19.513 | 1.00 | 16.73 | A | C |
| ATOM | 1661 | CD1 | LEU A 231 | 64.094 | 17.277 | 18.822 | 1.00 | 16.37 | A | C |
| ATOM | 1662 | CD2 | LEU A 231 | 66.353 | 16.285 | 18.498 | 1.00 | 16.99 | A | C |
| ATOM | 1663 | C | LEU A 231 | 64.647 | 17.582 | 22.515 | 1.00 | 22.71 | A | C |
| ATOM | 1664 | O | LEU A 231 | 65.340 | 16.619 | 22.813 | 1.00 | 24.85 | A | O |
| ATOM | 1665 | N | VAL A 232 | 63.516 | 17.898 | 23.130 | 1.00 | 23.84 | A | N |
| ATOM | 1666 | CA | VAL A 232 | 62.967 | 17.165 | 24.263 | 1.00 | 23.85 | A | C |
| ATOM | 1667 | CB | VAL A 232 | 61.820 | 17.972 | 24.892 | 1.00 | 24.09 | A | C |
| ATOM | 1668 | CG1 | VAL A 232 | 61.448 | 17.394 | 26.255 | 1.00 | 23.17 | A | C |
| ATOM | 1669 | CG2 | VAL A 232 | 62.222 | 19.431 | 24.987 | 1.00 | 21.20 | A | C |
| ATOM | 1670 | C | VAL A 232 | 62.447 | 15.777 | 23.927 | 1.00 | 25.57 | A | C |
| ATOM | 1671 | O | VAL A 232 | 61.821 | 15.576 | 22.887 | 1.00 | 25.74 | A | O |
| ATOM | 1672 | N | GLY A 233 | 62.692 | 14.828 | 24.829 | 1.00 | 26.04 | A | N |
| ATOM | 1673 | CA | GLY A 233 | 62.240 | 13.461 | 24.619 | 1.00 | 26.94 | A | C |
| ATOM | 1674 | C | GLY A 233 | 60.923 | 13.174 | 25.315 | 1.00 | 25.88 | A | C |
| ATOM | 1675 | O | GLY A 233 | 59.947 | 13.896 | 25.129 | 1.00 | 27.25 | A | O |
| ATOM | 1676 | N | GLY A 234 | 60.911 | 12.127 | 26.136 | 1.00 | 24.68 | A | N |
| ATOM | 1677 | CA | GLY A 234 | 59.709 | 11.734 | 26.852 | 1.00 | 23.13 | A | C |
| ATOM | 1678 | C | GLY A 234 | 58.901 | 12.858 | 27.480 | 1.00 | 22.51 | A | C |
| ATOM | 1679 | O | GLY A 234 | 57.669 | 12.808 | 27.483 | 1.00 | 21.53 | A | O |
| ATOM | 1680 | N | ALA A 235 | 59.587 | 13.864 | 28.016 | 1.00 | 20.33 | A | N |
| ATOM | 1681 | CA | ALA A 235 | 58.913 | 14.982 | 28.658 | 1.00 | 20.17 | A | C |
| ATOM | 1682 | CB | ALA A 235 | 59.937 | 15.975 | 29.209 | 1.00 | 21.87 | A | C |
| ATOM | 1683 | C | ALA A 235 | 57.968 | 15.685 | 27.696 | 1.00 | 20.55 | A | C |
| ATOM | 1684 | O | ALA A 235 | 57.055 | 16.380 | 28.133 | 1.00 | 21.17 | A | O |
| ATOM | 1685 | N | SER A 236 | 58.179 | 15.511 | 26.392 | 1.00 | 19.15 | A | N |
| ATOM | 1686 | CA | SER A 236 | 57.303 | 16.151 | 25.423 | 1.00 | 19.11 | A | C |
| ATOM | 1687 | CB | SER A 236 | 57.986 | 16.276 | 24.061 | 1.00 | 18.61 | A | C |
| ATOM | 1688 | OG | SER A 236 | 58.023 | 15.044 | 23.371 | 1.00 | 22.28 | A | O |
| ATOM | 1689 | C | SER A 236 | 55.993 | 15.373 | 25.285 | 1.00 | 20.21 | A | C |
| ATOM | 1690 | O | SER A 236 | 55.124 | 15.745 | 24.496 | 1.00 | 19.31 | A | O |
| ATOM | 1691 | N | LEU A 237 | 55.861 | 14.300 | 26.064 | 1.00 | 19.77 | A | N |
| ATOM | 1692 | CA | LEU A 237 | 54.657 | 13.470 | 26.063 | 1.00 | 21.74 | A | C |
| ATOM | 1693 | CB | LEU A 237 | 55.011 | 12.018 | 26.401 | 1.00 | 21.73 | A | C |
| ATOM | 1694 | CG | LEU A 237 | 55.272 | 11.064 | 25.227 | 1.00 | 24.24 | A | C |
| ATOM | 1695 | CD1 | LEU A 237 | 55.669 | 11.841 | 23.980 | 1.00 | 24.31 | A | C |
| ATOM | 1696 | CD2 | LEU A 237 | 56.354 | 10.064 | 25.620 | 1.00 | 23.37 | A | C |
| ATOM | 1697 | C | LEU A 237 | 53.591 | 13.977 | 27.034 | 1.00 | 21.75 | A | C |
| ATOM | 1698 | O | LEU A 237 | 52.479 | 13.461 | 27.060 | 1.00 | 22.31 | A | O |
| ATOM | 1699 | N | ASN A 238 | 53.939 | 14.972 | 27.844 | 1.00 | 22.17 | A | N |
| ATOM | 1700 | CA | ASN A 238 | 52.997 | 15.557 | 28.796 | 1.00 | 22.68 | A | C |
| ATOM | 1701 | CB | ASN A 238 | 53.491 | 15.358 | 30.234 | 1.00 | 21.33 | A | C |
| ATOM | 1702 | CG | ASN A 238 | 52.523 | 15.910 | 31.270 | 1.00 | 22.65 | A | C |
| ATOM | 1703 | OD1 | ASN A 238 | 52.555 | 17.091 | 31.596 | 1.00 | 22.15 | A | O |
| ATOM | 1704 | ND2 | ASN A 238 | 51.649 | 15.051 | 31.782 | 1.00 | 21.80 | A | N |
| ATOM | 1705 | C | ASN A 238 | 52.875 | 17.045 | 28.466 | 1.00 | 22.64 | A | C |
| ATOM | 1706 | O | ASN A 238 | 53.860 | 17.779 | 28.513 | 1.00 | 22.21 | A | O |
| ATOM | 1707 | N | ALA A 239 | 51.662 | 17.481 | 28.130 | 1.00 | 23.15 | A | N |
| ATOM | 1708 | CA | ALA A 239 | 51.422 | 18.869 | 27.745 | 1.00 | 23.18 | A | C |
| ATOM | 1709 | CB | ALA A 239 | 49.920 | 19.127 | 27.590 | 1.00 | 24.02 | A | C |
| ATOM | 1710 | C | ALA A 239 | 52.025 | 19.905 | 28.674 | 1.00 | 22.58 | A | C |
| ATOM | 1711 | O | ALA A 239 | 52.894 | 20.671 | 28.261 | 1.00 | 21.87 | A | O |
| ATOM | 1712 | N | ASP A 240 | 51.569 | 19.935 | 29.924 | 1.00 | 23.40 | A | N |
| ATOM | 1713 | CA | ASP A 240 | 52.065 | 20.924 | 30.885 | 1.00 | 24.99 | A | C |

FIGURE 9 - 26

36/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|--------|--------|------|-------|---|---|
| ATOM | 1714 | CB | ASP | A | 240 | 51.301 | 20.820 | 32.218 | 1.00 | 28.18 | A | C |
| ATOM | 1715 | CG | ASP | A | 240 | 49.787 | 21.046 | 32.058 | 1.00 | 34.40 | A | C |
| ATOM | 1716 | OD1 | ASP | A | 240 | 49.375 | 21.956 | 31.294 | 1.00 | 34.40 | A | O |
| ATOM | 1717 | OD2 | ASP | A | 240 | 49.003 | 20.319 | 32.710 | 1.00 | 35.74 | A | O |
| ATOM | 1718 | C | ASP | A | 240 | 53.583 | 20.873 | 31.139 | 1.00 | 23.24 | A | C |
| ATOM | 1719 | O | ASP | A | 240 | 54.215 | 21.914 | 31.296 | 1.00 | 22.53 | A | O |
| ATOM | 1720 | N | GLU | A | 241 | 54.168 | 19.676 | 31.172 | 1.00 | 22.93 | A | N |
| ATOM | 1721 | CA | GLU | A | 241 | 55.610 | 19.558 | 31.399 | 1.00 | 23.44 | A | C |
| ATOM | 1722 | CB | GLU | A | 241 | 56.019 | 18.087 | 31.594 | 1.00 | 23.75 | A | C |
| ATOM | 1723 | CG | GLU | A | 241 | 57.496 | 17.884 | 31.983 | 1.00 | 23.96 | A | C |
| ATOM | 1724 | CD | GLU | A | 241 | 57.846 | 16.428 | 32.310 | 1.00 | 25.18 | A | C |
| ATOM | 1725 | OE1 | GLU | A | 241 | 57.309 | 15.511 | 31.658 | 1.00 | 24.74 | A | O |
| ATOM | 1726 | OE2 | GLU | A | 241 | 58.673 | 16.197 | 33.211 | 1.00 | 26.97 | A | O |
| ATOM | 1727 | C | GLU | A | 241 | 56.362 | 20.162 | 30.209 | 1.00 | 21.74 | A | C |
| ATOM | 1728 | O | GLU | A | 241 | 57.387 | 20.813 | 30.382 | 1.00 | 21.93 | A | O |
| ATOM | 1729 | N | PHE | A | 242 | 55.846 | 19.935 | 29.005 | 1.00 | 21.24 | A | N |
| ATOM | 1730 | CA | PHE | A | 242 | 56.457 | 20.468 | 27.791 | 1.00 | 20.79 | A | C |
| ATOM | 1731 | CB | PHE | A | 242 | 55.753 | 19.906 | 26.554 | 1.00 | 18.97 | A | C |
| ATOM | 1732 | CG | PHE | A | 242 | 56.376 | 20.329 | 25.258 | 1.00 | 16.57 | A | C |
| ATOM | 1733 | CD1 | PHE | A | 242 | 57.694 | 20.003 | 24.966 | 1.00 | 17.22 | A | C |
| ATOM | 1734 | CD2 | PHE | A | 242 | 55.645 | 21.055 | 24.325 | 1.00 | 17.37 | A | C |
| ATOM | 1735 | CE1 | PHE | A | 242 | 58.267 | 20.389 | 23.776 | 1.00 | 12.97 | A | C |
| ATOM | 1736 | CE2 | PHE | A | 242 | 56.214 | 21.446 | 23.130 | 1.00 | 14.33 | A | C |
| ATOM | 1737 | CZ | PHE | A | 242 | 57.529 | 21.111 | 22.857 | 1.00 | 16.37 | A | C |
| ATOM | 1738 | C | PHE | A | 242 | 56.339 | 21.990 | 27.797 | 1.00 | 21.03 | A | C |
| ATOM | 1739 | O | PHE | A | 242 | 57.229 | 22.697 | 27.332 | 1.00 | 21.95 | A | O |
| ATOM | 1740 | N | GLY | A | 243 | 55.228 | 22.484 | 28.330 | 1.00 | 20.92 | A | N |
| ATOM | 1741 | CA | GLY | A | 243 | 55.015 | 23.915 | 28.401 | 1.00 | 19.75 | A | C |
| ATOM | 1742 | C | GLY | A | 243 | 56.014 | 24.589 | 29.324 | 1.00 | 20.31 | A | C |
| ATOM | 1743 | O | GLY | A | 243 | 56.532 | 25.667 | 29.025 | 1.00 | 18.71 | A | O |
| ATOM | 1744 | N | ALA | A | 244 | 56.295 | 23.953 | 30.455 | 1.00 | 20.28 | A | N |
| ATOM | 1745 | CA | ALA | A | 244 | 57.242 | 24.521 | 31.400 | 1.00 | 22.47 | A | C |
| ATOM | 1746 | CB | ALA | A | 244 | 57.265 | 23.700 | 32.674 | 1.00 | 22.84 | A | C |
| ATOM | 1747 | C | ALA | A | 244 | 58.636 | 24.588 | 30.785 | 1.00 | 23.43 | A | C |
| ATOM | 1748 | O | ALA | A | 244 | 59.420 | 25.479 | 31.107 | 1.00 | 24.58 | A | O |
| ATOM | 1749 | N | ILE | A | 245 | 58.950 | 23.650 | 29.895 | 1.00 | 24.41 | A | N |
| ATOM | 1750 | CA | ILE | A | 245 | 60.269 | 23.645 | 29.259 | 1.00 | 23.65 | A | C |
| ATOM | 1751 | CB | ILE | A | 245 | 60.573 | 22.282 | 28.618 | 1.00 | 22.77 | A | C |
| ATOM | 1752 | CG2 | ILE | A | 245 | 61.906 | 22.337 | 27.880 | 1.00 | 23.46 | A | C |
| ATOM | 1753 | CG1 | ILE | A | 245 | 60.580 | 21.206 | 29.709 | 1.00 | 23.75 | A | C |
| ATOM | 1754 | CD1 | ILE | A | 245 | 60.959 | 19.830 | 29.235 | 1.00 | 24.48 | A | C |
| ATOM | 1755 | C | ILE | A | 245 | 60.374 | 24.749 | 28.207 | 1.00 | 23.26 | A | C |
| ATOM | 1756 | O | ILE | A | 245 | 61.423 | 25.379 | 28.069 | 1.00 | 22.75 | A | O |
| ATOM | 1757 | N | CYS | A | 246 | 59.289 | 24.977 | 27.469 | 1.00 | 22.68 | A | N |
| ATOM | 1758 | CA | CYS | A | 246 | 59.269 | 26.029 | 26.458 | 1.00 | 24.31 | A | C |
| ATOM | 1759 | CB | CYS | A | 246 | 57.934 | 26.042 | 25.707 | 1.00 | 23.73 | A | C |
| ATOM | 1760 | SG | CYS | A | 246 | 57.624 | 24.624 | 24.638 | 1.00 | 24.50 | A | S |
| ATOM | 1761 | C | CYS | A | 246 | 59.446 | 27.368 | 27.166 | 1.00 | 24.44 | A | C |
| ATOM | 1762 | O | CYS | A | 246 | 60.242 | 28.213 | 26.757 | 1.00 | 26.55 | A | O |
| ATOM | 1763 | N | ARG | A | 247 | 58.695 | 27.549 | 28.241 | 1.00 | 24.08 | A | N |
| ATOM | 1764 | CA | ARG | A | 247 | 58.750 | 28.787 | 29.003 | 1.00 | 26.08 | A | C |
| ATOM | 1765 | CB | ARG | A | 247 | 57.763 | 28.704 | 30.171 | 1.00 | 27.21 | A | C |
| ATOM | 1766 | CG | ARG | A | 247 | 57.135 | 30.025 | 30.564 | 1.00 | 30.39 | A | C |
| ATOM | 1767 | CD | ARG | A | 247 | 55.931 | 29.800 | 31.467 | 1.00 | 31.58 | A | C |
| ATOM | 1768 | NE | ARG | A | 247 | 54.775 | 29.253 | 30.747 | 1.00 | 30.93 | A | N |
| ATOM | 1769 | CZ | ARG | A | 247 | 54.268 | 28.035 | 30.940 | 1.00 | 30.00 | A | C |
| ATOM | 1770 | NH1 | ARG | A | 247 | 54.815 | 27.211 | 31.832 | 1.00 | 27.30 | A | N |
| ATOM | 1771 | NH2 | ARG | A | 247 | 53.183 | 27.653 | 30.262 | 1.00 | 29.94 | A | N |
| ATOM | 1772 | C | ARG | A | 247 | 60.170 | 29.048 | 29.506 | 1.00 | 25.08 | A | C |
| ATOM | 1773 | O | ARG | A | 247 | 60.642 | 30.183 | 29.495 | 1.00 | 24.80 | A | O |
| ATOM | 1774 | N | ALA | A | 248 | 60.854 | 27.988 | 29.926 | 1.00 | 26.03 | A | N |
| ATOM | 1775 | CA | ALA | A | 248 | 62.221 | 28.107 | 30.427 | 1.00 | 25.39 | A | C |
| ATOM | 1776 | CB | ALA | A | 248 | 62.715 | 26.756 | 30.914 | 1.00 | 26.95 | A | C |
| ATOM | 1777 | C | ALA | A | 248 | 63.167 | 28.649 | 29.365 | 1.00 | 25.92 | A | C |
| ATOM | 1778 | O | ALA | A | 248 | 64.271 | 29.085 | 29.667 | 1.00 | 26.97 | A | O |
| ATOM | 1779 | N | ALA | A | 249 | 62.738 | 28.620 | 28.113 | 1.00 | 27.25 | A | N |

FIGURE 9 - 27

SUBSTITUTE SHEET (RULE 26)

37/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|--------|------|-------|---|---|
| ATOM | 1780 | CA | ALA | A | 249 | 63.579 | 29.121 | 27.035 | 1.00 | 28.40 | A | C |
| ATOM | 1781 | CB | ALA | A | 249 | 63.377 | 28.274 | 25.791 | 1.00 | 26.39 | A | C |
| ATOM | 1782 | C | ALA | A | 249 | 63.301 | 30.592 | 26.714 | 1.00 | 29.71 | A | C |
| ATOM | 1783 | O | ALA | A | 249 | 64.059 | 31.209 | 25.971 | 1.00 | 30.25 | A | O |
| ATOM | 1784 | N | GLY | A | 250 | 62.224 | 31.146 | 27.271 | 1.00 | 32.16 | A | N |
| ATOM | 1785 | CA | GLY | A | 250 | 61.873 | 32.535 | 26.999 | 1.00 | 36.17 | A | C |
| ATOM | 1786 | C | GLY | A | 250 | 62.337 | 33.560 | 28.022 | 1.00 | 39.60 | A | C |
| ATOM | 1787 | O | GLY | A | 250 | 63.221 | 34.399 | 27.693 | 1.00 | 41.59 | A | O |
| ATOM | 1788 | OXT | GLY | A | 250 | 61.822 | 33.535 | 29.161 | 1.00 | 42.10 | A | O |
| TER | 1789 | | GLY | A | 250 | | | | | | A | |
| ATOM | 1790 | CB | MET | B | 1 | 38.254 | -23.711 | -0.133 | 1.00 | 39.29 | B | C |
| ATOM | 1791 | CG | MET | B | 1 | 37.707 | -24.999 | -0.760 | 1.00 | 41.51 | B | C |
| ATOM | 1792 | SD | MET | B | 1 | 38.988 | -25.999 | -1.568 | 1.00 | 43.23 | B | S |
| ATOM | 1793 | CE | MET | B | 1 | 38.759 | -27.639 | -0.702 | 1.00 | 41.88 | B | C |
| ATOM | 1794 | C | MET | B | 1 | 40.027 | -22.521 | 1.143 | 1.00 | 38.04 | B | C |
| ATOM | 1795 | O | MET | B | 1 | 39.397 | -21.686 | 1.796 | 1.00 | 37.73 | B | O |
| ATOM | 1796 | N | MET | B | 1 | 38.926 | -24.565 | 2.117 | 1.00 | 37.65 | B | N |
| ATOM | 1797 | CA | MET | B | 1 | 39.412 | -23.895 | 0.860 | 1.00 | 38.46 | B | C |
| ATOM | 1798 | N | ARG | B | 2 | 41.245 | -22.286 | 0.648 | 1.00 | 36.65 | B | N |
| ATOM | 1799 | CA | ARG | B | 2 | 41.916 | -20.996 | 0.855 | 1.00 | 34.99 | B | C |
| ATOM | 1800 | CB | ARG | B | 2 | 43.265 | -20.940 | 0.127 | 1.00 | 34.38 | B | C |
| ATOM | 1801 | CG | ARG | B | 2 | 44.359 | -21.824 | 0.705 | 1.00 | 33.70 | B | C |
| ATOM | 1802 | CD | ARG | B | 2 | 45.687 | -21.565 | -0.003 | 1.00 | 30.29 | B | C |
| ATOM | 1803 | NE | ARG | B | 2 | 45.641 | -21.950 | -1.413 | 1.00 | 28.14 | B | N |
| ATOM | 1804 | CZ | ARG | B | 2 | 45.937 | -23.163 | -1.873 | 1.00 | 29.01 | B | C |
| ATOM | 1805 | NH1 | ARG | B | 2 | 46.312 | -24.116 | -1.038 | 1.00 | 26.43 | B | N |
| ATOM | 1806 | NH2 | ARG | B | 2 | 45.844 | -23.432 | -3.174 | 1.00 | 29.39 | B | N |
| ATOM | 1807 | C | ARG | B | 2 | 41.038 | -19.883 | 0.311 | 1.00 | 33.64 | B | C |
| ATOM | 1808 | O | ARG | B | 2 | 40.539 | -19.969 | -0.810 | 1.00 | 34.35 | B | O |
| ATOM | 1809 | N | ARG | B | 3 | 40.864 | -18.830 | 1.095 | 1.00 | 32.19 | B | N |
| ATOM | 1810 | CA | ARG | B | 3 | 40.038 | -17.705 | 0.676 | 1.00 | 29.95 | B | C |
| ATOM | 1811 | CB | ARG | B | 3 | 39.452 | -17.055 | 1.927 | 1.00 | 30.86 | B | C |
| ATOM | 1812 | CG | ARG | B | 3 | 38.335 | -16.067 | 1.698 | 1.00 | 32.31 | B | C |
| ATOM | 1813 | CD | ARG | B | 3 | 37.760 | -15.587 | 3.041 | 1.00 | 33.83 | B | C |
| ATOM | 1814 | NE | ARG | B | 3 | 36.913 | -14.410 | 2.856 | 1.00 | 35.04 | B | N |
| ATOM | 1815 | CZ | ARG | B | 3 | 35.695 | -14.438 | 2.328 | 1.00 | 35.89 | B | C |
| ATOM | 1816 | NH1 | ARG | B | 3 | 35.166 | -15.593 | 1.941 | 1.00 | 36.43 | B | N |
| ATOM | 1817 | NH2 | ARG | B | 3 | 35.018 | -13.305 | 2.158 | 1.00 | 36.37 | B | N |
| ATOM | 1818 | C | ARG | B | 3 | 40.888 | -16.708 | -0.121 | 1.00 | 29.77 | B | C |
| ATOM | 1819 | O | ARG | B | 3 | 41.971 | -16.327 | 0.316 | 1.00 | 30.29 | B | O |
| ATOM | 1820 | N | PRO | B | 4 | 40.423 | -16.284 | -1.313 | 1.00 | 29.13 | B | N |
| ATOM | 1821 | CD | PRO | B | 4 | 39.280 | -16.795 | -2.090 | 1.00 | 29.44 | B | C |
| ATOM | 1822 | CA | PRO | B | 4 | 41.204 | -15.327 | -2.112 | 1.00 | 28.31 | B | C |
| ATOM | 1823 | CB | PRO | B | 4 | 40.353 | -15.155 | -3.368 | 1.00 | 29.26 | B | C |
| ATOM | 1824 | CG | PRO | B | 4 | 39.702 | -16.496 | -3.517 | 1.00 | 29.38 | B | C |
| ATOM | 1825 | C | PRO | B | 4 | 41.465 | -13.986 | -1.402 | 1.00 | 29.09 | B | C |
| ATOM | 1826 | O | PRO | B | 4 | 40.630 | -13.503 | -0.626 | 1.00 | 26.82 | B | O |
| ATOM | 1827 | N | LEU | B | 5 | 42.621 | -13.386 | -1.688 | 1.00 | 27.53 | B | N |
| ATOM | 1828 | CA | LEU | B | 5 | 43.001 | -12.107 | -1.085 | 1.00 | 26.45 | B | C |
| ATOM | 1829 | CB | LEU | B | 5 | 44.046 | -12.316 | 0.018 | 1.00 | 25.53 | B | C |
| ATOM | 1830 | CG | LEU | B | 5 | 44.799 | -11.068 | 0.510 | 1.00 | 24.95 | B | C |
| ATOM | 1831 | CD1 | LEU | B | 5 | 43.837 | -10.111 | 1.177 | 1.00 | 25.86 | B | C |
| ATOM | 1832 | CD2 | LEU | B | 5 | 45.893 | -11.466 | 1.482 | 1.00 | 23.55 | B | C |
| ATOM | 1833 | C | LEU | B | 5 | 43.572 | -11.153 | -2.120 | 1.00 | 26.71 | B | C |
| ATOM | 1834 | O | LEU | B | 5 | 44.498 | -11.497 | -2.854 | 1.00 | 27.09 | B | O |
| ATOM | 1835 | N | VAL | B | 6 | 43.015 | -9.950 | -2.178 | 1.00 | 26.99 | B | N |
| ATOM | 1836 | CA | VAL | B | 6 | 43.495 | -8.939 | -3.106 | 1.00 | 27.35 | B | C |
| ATOM | 1837 | CB | VAL | B | 6 | 42.406 | -8.533 | -4.101 | 1.00 | 27.88 | B | C |
| ATOM | 1838 | CG1 | VAL | B | 6 | 42.920 | -7.413 | -4.999 | 1.00 | 26.82 | B | C |
| ATOM | 1839 | CG2 | VAL | B | 6 | 41.987 | -9.749 | -4.920 | 1.00 | 28.26 | B | C |
| ATOM | 1840 | C | VAL | B | 6 | 43.898 | -7.727 | -2.283 | 1.00 | 28.26 | B | C |
| ATOM | 1841 | O | VAL | B | 6 | 43.043 | -7.045 | -1.721 | 1.00 | 30.19 | B | O |
| ATOM | 1842 | N | ALA | B | 7 | 45.199 | -7.461 | -2.209 | 1.00 | 28.32 | B | N |
| ATOM | 1843 | CA | ALA | B | 7 | 45.709 | -6.341 | -1.422 | 1.00 | 27.60 | B | C |
| ATOM | 1844 | CB | ALA | B | 7 | 46.636 | -6.863 | -0.348 | 1.00 | 27.24 | B | C |
| ATOM | 1845 | C | ALA | B | 7 | 46.432 | -5.305 | -2.279 | 1.00 | 27.32 | B | C |

FIGURE 9 - 28

38/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 1846 | O | ALA | B | 7 | 47.255 | -5.644 | -3.130 | 1.00 | 27.65 | B | O |
| ATOM | 1847 | N | GLY | B | 8 | 46.130 | -4.036 | -2.029 | 1.00 | 26.92 | B | N |
| ATOM | 1848 | CA | GLY | B | 8 | 46.737 | -2.965 | -2.797 | 1.00 | 26.72 | B | C |
| ATOM | 1849 | C | GLY | B | 8 | 47.813 | -2.189 | -2.062 | 1.00 | 26.89 | B | C |
| ATOM | 1850 | O | GLY | B | 8 | 47.592 | -1.667 | -0.972 | 1.00 | 26.02 | B | O |
| ATOM | 1851 | N | ASN | B | 9 | 48.991 | -2.120 | -2.670 | 1.00 | 26.08 | B | N |
| ATOM | 1852 | CA | ASN | B | 9 | 50.115 | -1.403 | -2.094 | 1.00 | 24.90 | B | C |
| ATOM | 1853 | CB | ASN | B | 9 | 51.409 | -2.180 | -2.313 | 1.00 | 23.23 | B | C |
| ATOM | 1854 | CG | ASN | B | 9 | 52.612 | -1.445 | -1.788 | 1.00 | 21.91 | B | C |
| ATOM | 1855 | OD1 | ASN | B | 9 | 52.478 | -0.526 | -0.989 | 1.00 | 22.31 | B | O |
| ATOM | 1856 | ND2 | ASN | B | 9 | 53.799 | -1.851 | -2.222 | 1.00 | 19.96 | B | N |
| ATOM | 1857 | C | ASN | B | 9 | 50.211 | -0.045 | -2.766 | 1.00 | 23.88 | B | C |
| ATOM | 1858 | O | ASN | B | 9 | 50.701 | 0.066 | -3.883 | 1.00 | 23.71 | B | O |
| ATOM | 1859 | N | TRP | B | 10 | 49.738 | 0.983 | -2.075 | 1.00 | 23.64 | B | N |
| ATOM | 1860 | CA | TRP | B | 10 | 49.748 | 2.335 | -2.611 | 1.00 | 23.50 | B | C |
| ATOM | 1861 | CB | TRP | B | 10 | 48.905 | 3.257 | -1.719 | 1.00 | 24.27 | B | C |
| ATOM | 1862 | CG | TRP | B | 10 | 47.450 | 2.907 | -1.635 | 1.00 | 23.37 | B | C |
| ATOM | 1863 | CD2 | TRP | B | 10 | 46.424 | 3.663 | -0.973 | 1.00 | 25.18 | B | C |
| ATOM | 1864 | CE2 | TRP | B | 10 | 45.212 | 2.950 | -1.123 | 1.00 | 23.26 | B | C |
| ATOM | 1865 | CE3 | TRP | B | 10 | 46.412 | 4.874 | -0.266 | 1.00 | 23.36 | B | C |
| ATOM | 1866 | CD1 | TRP | B | 10 | 46.833 | 1.796 | -2.149 | 1.00 | 25.60 | B | C |
| ATOM | 1867 | NE1 | TRP | B | 10 | 45.486 | 1.816 | -1.844 | 1.00 | 24.60 | B | N |
| ATOM | 1868 | CE2 | TRP | B | 10 | 44.001 | 3.410 | -0.588 | 1.00 | 24.13 | B | C |
| ATOM | 1869 | CE3 | TRP | B | 10 | 45.207 | 5.329 | 0.265 | 1.00 | 22.31 | B | C |
| ATOM | 1870 | CH2 | TRP | B | 10 | 44.022 | 4.597 | 0.100 | 1.00 | 23.37 | B | C |
| ATOM | 1871 | C | TRP | B | 10 | 51.158 | 2.901 | -2.720 | 1.00 | 23.33 | B | C |
| ATOM | 1872 | O | TRP | B | 10 | 51.393 | 3.859 | -3.459 | 1.00 | 26.44 | B | O |
| ATOM | 1873 | N | LYS | B | 11 | 52.090 | 2.301 | -1.989 | 1.00 | 22.51 | B | N |
| ATOM | 1874 | CA | LYS | B | 11 | 53.476 | 2.763 | -1.961 | 1.00 | 24.92 | B | C |
| ATOM | 1875 | CB | LYS | B | 11 | 54.137 | 2.588 | -3.336 | 1.00 | 24.05 | B | C |
| ATOM | 1876 | CG | LYS | B | 11 | 54.235 | 1.115 | -3.764 | 1.00 | 22.53 | B | C |
| ATOM | 1877 | CD | LYS | B | 11 | 55.080 | 0.916 | -5.023 | 1.00 | 23.18 | B | C |
| ATOM | 1878 | CE | LYS | B | 11 | 55.147 | -0.561 | -5.391 | 1.00 | 23.09 | B | C |
| ATOM | 1879 | NZ | LYS | B | 11 | 56.049 | -0.856 | -6.533 | 1.00 | 24.44 | B | N |
| ATOM | 1880 | C | LYS | B | 11 | 53.570 | 4.220 | -1.468 | 1.00 | 24.17 | B | C |
| ATOM | 1881 | O | LYS | B | 11 | 52.679 | 4.684 | -0.763 | 1.00 | 23.95 | B | O |
| ATOM | 1882 | N | MET | B | 12 | 54.637 | 4.931 | -1.831 | 1.00 | 24.23 | B | N |
| ATOM | 1883 | CA | MET | B | 12 | 54.843 | 6.321 | -1.376 | 1.00 | 23.99 | B | C |
| ATOM | 1884 | CB | MET | B | 12 | 56.356 | 6.657 | -1.450 | 1.00 | 20.54 | B | C |
| ATOM | 1885 | CG | MET | B | 12 | 56.822 | 7.931 | -0.704 | 1.00 | 17.88 | B | C |
| ATOM | 1886 | SD | MET | B | 12 | 58.635 | 8.096 | -0.443 | 1.00 | 8.79 | B | S |
| ATOM | 1887 | CE | MET | B | 12 | 59.199 | 8.336 | -2.054 | 1.00 | 14.14 | B | C |
| ATOM | 1888 | C | MET | B | 12 | 54.004 | 7.341 | -2.169 | 1.00 | 22.63 | B | C |
| ATOM | 1889 | O | MET | B | 12 | 54.551 | 8.218 | -2.829 | 1.00 | 23.70 | B | O |
| ATOM | 1890 | N | HIS | B | 13 | 52.677 | 7.232 | -2.086 | 1.00 | 23.59 | B | N |
| ATOM | 1891 | CA | HIS | B | 13 | 51.783 | 8.135 | -2.823 | 1.00 | 24.71 | B | C |
| ATOM | 1892 | CB | HIS | B | 13 | 51.226 | 7.461 | -4.075 | 1.00 | 24.25 | B | C |
| ATOM | 1893 | CG | HIS | B | 13 | 52.249 | 7.148 | -5.115 | 1.00 | 25.34 | B | C |
| ATOM | 1894 | CD2 | HIS | B | 13 | 52.654 | 7.847 | -6.200 | 1.00 | 24.96 | B | C |
| ATOM | 1895 | ND1 | HIS | B | 13 | 52.959 | 5.966 | -5.130 | 1.00 | 26.90 | B | N |
| ATOM | 1896 | CE1 | HIS | B | 13 | 53.755 | 5.950 | -6.185 | 1.00 | 27.01 | B | C |
| ATOM | 1897 | NE2 | HIS | B | 13 | 53.589 | 7.079 | -6.850 | 1.00 | 26.99 | B | N |
| ATOM | 1898 | C | HIS | B | 13 | 50.575 | 8.686 | -2.078 | 1.00 | 25.92 | B | C |
| ATOM | 1899 | O | HIS | B | 13 | 50.131 | 8.131 | -1.073 | 1.00 | 26.14 | B | O |
| ATOM | 1900 | N | GLY | B | 14 | 50.033 | 9.777 | -2.616 | 1.00 | 26.76 | B | N |
| ATOM | 1901 | CA | GLY | B | 14 | 48.849 | 10.390 | -2.048 | 1.00 | 26.86 | B | C |
| ATOM | 1902 | C | GLY | B | 14 | 48.996 | 11.700 | -1.293 | 1.00 | 28.49 | B | C |
| ATOM | 1903 | O | GLY | B | 14 | 50.073 | 12.043 | -0.805 | 1.00 | 28.16 | B | O |
| ATOM | 1904 | N | THR | B | 15 | 47.881 | 12.427 | -1.231 | 1.00 | 28.51 | B | N |
| ATOM | 1905 | CA | THR | B | 15 | 47.741 | 13.693 | -0.522 | 1.00 | 30.91 | B | C |
| ATOM | 1906 | CB | THR | B | 15 | 47.760 | 14.898 | -1.490 | 1.00 | 31.69 | B | C |
| ATOM | 1907 | OG1 | THR | B | 15 | 46.666 | 14.789 | -2.408 | 1.00 | 33.09 | B | O |
| ATOM | 1908 | CG2 | THR | B | 15 | 49.075 | 14.942 | -2.282 | 1.00 | 31.05 | B | C |
| ATOM | 1909 | C | THR | B | 15 | 46.348 | 13.570 | 0.118 | 1.00 | 32.25 | B | C |
| ATOM | 1910 | O | THR | B | 15 | 45.609 | 12.649 | -0.222 | 1.00 | 32.77 | B | O |
| ATOM | 1911 | N | HIS | B | 16 | 45.987 | 14.458 | 1.042 | 1.00 | 31.79 | B | N |

FIGURE 9 - 29

SUBSTITUTE SHEET (RULE 26)

39/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 1912 | CA | HIS | B | 16 | 44.667 | 14.377 | 1.680 | 1.00 | 34.07 | B | C |
| ATOM | 1913 | CB | HIS | B | 16 | 44.413 | 15.582 | 2.607 | 1.00 | 36.99 | B | C |
| ATOM | 1914 | CG | HIS | B | 16 | 44.949 | 15.411 | 3.996 | 1.00 | 41.23 | B | C |
| ATOM | 1915 | CD2 | HIS | B | 16 | 45.811 | 16.167 | 4.720 | 1.00 | 43.07 | B | C |
| ATOM | 1916 | ND1 | HIS | B | 16 | 44.581 | 14.363 | 4.813 | 1.00 | 43.20 | B | N |
| ATOM | 1917 | CE1 | HIS | B | 16 | 45.194 | 14.479 | 5.979 | 1.00 | 44.48 | B | C |
| ATOM | 1918 | NE2 | HIS | B | 16 | 45.947 | 15.567 | 5.949 | 1.00 | 44.04 | B | N |
| ATOM | 1919 | C | HIS | B | 16 | 43.536 | 14.318 | 0.658 | 1.00 | 32.73 | B | C |
| ATOM | 1920 | O | HIS | B | 16 | 42.553 | 13.600 | 0.844 | 1.00 | 32.91 | B | O |
| ATOM | 1921 | N | SER | B | 17 | 43.686 | 15.084 | -0.414 | 1.00 | 32.54 | B | N |
| ATOM | 1922 | CA | SER | B | 17 | 42.683 | 15.158 | -1.466 | 1.00 | 33.11 | B | C |
| ATOM | 1923 | CB | SER | B | 17 | 42.908 | 16.407 | -2.304 | 1.00 | 33.96 | B | C |
| ATOM | 1924 | OG | SER | B | 17 | 42.091 | 16.365 | -3.456 | 1.00 | 37.21 | B | O |
| ATOM | 1925 | C | SER | B | 17 | 42.614 | 13.946 | -2.390 | 1.00 | 32.16 | B | C |
| ATOM | 1926 | O | SER | B | 17 | 41.523 | 13.475 | -2.696 | 1.00 | 32.55 | B | O |
| ATOM | 1927 | N | SER | B | 18 | 43.762 | 13.456 | -2.852 | 1.00 | 30.83 | B | N |
| ATOM | 1928 | CA | SER | B | 18 | 43.782 | 12.286 | -3.736 | 1.00 | 30.50 | B | C |
| ATOM | 1929 | CB | SER | B | 18 | 45.195 | 12.033 | -4.273 | 1.00 | 31.03 | B | C |
| ATOM | 1930 | OG | SER | B | 18 | 46.046 | 11.526 | -3.260 | 1.00 | 28.73 | B | O |
| ATOM | 1931 | C | SER | B | 18 | 43.298 | 11.035 | -2.994 | 1.00 | 30.38 | B | C |
| ATOM | 1932 | O | SER | B | 18 | 42.602 | 10.196 | -3.562 | 1.00 | 30.63 | B | O |
| ATOM | 1933 | N | VAL | B | 19 | 43.672 | 10.915 | -1.723 | 1.00 | 29.38 | B | N |
| ATOM | 1934 | CA | VAL | B | 19 | 43.265 | 9.774 | -0.915 | 1.00 | 29.67 | B | C |
| ATOM | 1935 | CB | VAL | B | 19 | 43.971 | 9.771 | 0.454 | 1.00 | 26.63 | B | C |
| ATOM | 1936 | CG1 | VAL | B | 19 | 43.428 | 8.639 | 1.309 | 1.00 | 25.40 | B | C |
| ATOM | 1937 | CG2 | VAL | B | 19 | 45.471 | 9.647 | 0.267 | 1.00 | 24.86 | B | C |
| ATOM | 1938 | C | VAL | B | 19 | 41.755 | 9.808 | -0.674 | 1.00 | 31.95 | B | C |
| ATOM | 1939 | O | VAL | B | 19 | 41.076 | 8.781 | -0.762 | 1.00 | 32.34 | B | O |
| ATOM | 1940 | N | ALA | B | 20 | 41.234 | 10.989 | -0.352 | 1.00 | 33.03 | B | N |
| ATOM | 1941 | CA | ALA | B | 20 | 39.809 | 11.127 | -0.107 | 1.00 | 35.20 | B | C |
| ATOM | 1942 | CB | ALA | B | 20 | 39.479 | 12.566 | 0.267 | 1.00 | 36.29 | B | C |
| ATOM | 1943 | C | ALA | B | 20 | 39.065 | 10.725 | -1.374 | 1.00 | 36.05 | B | C |
| ATOM | 1944 | O | ALA | B | 20 | 37.976 | 10.152 | -1.318 | 1.00 | 35.79 | B | O |
| ATOM | 1945 | N | GLU | B | 21 | 39.671 | 11.021 | -2.519 | 1.00 | 37.66 | B | N |
| ATOM | 1946 | CA | GLU | B | 21 | 39.073 | 10.693 | -3.810 | 1.00 | 39.08 | B | C |
| ATOM | 1947 | CB | GLU | B | 21 | 39.830 | 11.406 | -4.936 | 1.00 | 42.14 | B | C |
| ATOM | 1948 | CG | GLU | B | 21 | 39.014 | 11.631 | -6.204 | 1.00 | 47.37 | B | C |
| ATOM | 1949 | CD | GLU | B | 21 | 38.055 | 12.818 | -6.080 | 1.00 | 51.11 | B | C |
| ATOM | 1950 | OE1 | GLU | B | 21 | 38.539 | 13.983 | -6.033 | 1.00 | 52.26 | B | O |
| ATOM | 1951 | OE2 | GLU | B | 21 | 36.820 | 12.586 | -6.028 | 1.00 | 51.43 | B | O |
| ATOM | 1952 | C | GLU | B | 21 | 39.134 | 9.174 | -4.011 | 1.00 | 37.78 | B | C |
| ATOM | 1953 | O | GLU | B | 21 | 38.166 | 8.560 | -4.457 | 1.00 | 37.86 | B | O |
| ATOM | 1954 | N | LEU | B | 22 | 40.267 | 8.567 | -3.675 | 1.00 | 35.56 | B | N |
| ATOM | 1955 | CA | LEU | B | 22 | 40.400 | 7.126 | -3.827 | 1.00 | 34.47 | B | C |
| ATOM | 1956 | CB | LEU | B | 22 | 41.842 | 6.680 | -3.540 | 1.00 | 33.30 | B | C |
| ATOM | 1957 | CG | LEU | B | 22 | 42.174 | 5.213 | -3.857 | 1.00 | 31.98 | B | C |
| ATOM | 1958 | CD1 | LEU | B | 22 | 41.720 | 4.860 | -5.264 | 1.00 | 29.19 | B | C |
| ATOM | 1959 | CD2 | LEU | B | 22 | 43.662 | 4.989 | -3.711 | 1.00 | 31.33 | B | C |
| ATOM | 1960 | C | LEU | B | 22 | 39.422 | 6.414 | -2.891 | 1.00 | 34.47 | B | C |
| ATOM | 1961 | O | LEU | B | 22 | 38.770 | 5.446 | -3.292 | 1.00 | 34.84 | B | O |
| ATOM | 1962 | N | ILE | B | 23 | 39.307 | 6.902 | -1.656 | 1.00 | 34.19 | B | N |
| ATOM | 1963 | CA | ILE | B | 23 | 38.388 | 6.315 | -0.678 | 1.00 | 36.16 | B | C |
| ATOM | 1964 | CB | ILE | B | 23 | 38.422 | 7.068 | 0.671 | 1.00 | 35.63 | B | C |
| ATOM | 1965 | CG2 | ILE | B | 23 | 37.284 | 6.585 | 1.561 | 1.00 | 33.94 | B | C |
| ATOM | 1966 | CG1 | ILE | B | 23 | 39.764 | 6.862 | 1.373 | 1.00 | 35.67 | B | C |
| ATOM | 1967 | CD1 | ILE | B | 23 | 39.860 | 7.600 | 2.702 | 1.00 | 34.50 | B | C |
| ATOM | 1968 | C | ILE | B | 23 | 36.949 | 6.389 | -1.194 | 1.00 | 37.68 | B | C |
| ATOM | 1969 | O | ILE | B | 23 | 36.133 | 5.503 | -0.946 | 1.00 | 36.61 | B | O |
| ATOM | 1970 | N | LYS | B | 24 | 36.645 | 7.468 | -1.900 | 1.00 | 39.87 | B | N |
| ATOM | 1971 | CA | LYS | B | 24 | 35.311 | 7.661 | -2.439 | 1.00 | 41.28 | B | C |
| ATOM | 1972 | CB | LYS | B | 24 | 35.257 | 8.982 | -3.210 | 1.00 | 42.72 | B | C |
| ATOM | 1973 | CG | LYS | B | 24 | 33.848 | 9.512 | -3.451 | 1.00 | 45.01 | B | C |
| ATOM | 1974 | CD | LYS | B | 24 | 33.892 | 10.865 | -4.142 | 1.00 | 45.93 | B | C |
| ATOM | 1975 | CE | LYS | B | 24 | 32.513 | 11.504 | -4.199 | 1.00 | 46.92 | B | C |
| ATOM | 1976 | NZ | LYS | B | 24 | 32.587 | 12.873 | -4.795 | 1.00 | 48.33 | B | N |
| ATOM | 1977 | C | LYS | B | 24 | 35.002 | 6.493 | -3.367 | 1.00 | 41.36 | B | C |

FIGURE 9 - 30

SUBSTITUTE SHEET (RULE 26)

40/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|----|--------|---------|--------|------|-------|---|---|
| ATOM | 1978 | O | LYS | B | 24 | 34.050 | 5.746 | -3.138 | 1.00 | 41.22 | B | O |
| ATOM | 1979 | N | GLY | B | 25 | 35.825 | 6.339 | -4.404 | 1.00 | 40.93 | B | N |
| ATOM | 1980 | CA | GLY | B | 25 | 35.639 | 5.267 | -5.368 | 1.00 | 41.06 | B | C |
| ATOM | 1981 | C | GLY | B | 25 | 35.613 | 3.879 | -4.757 | 1.00 | 40.74 | B | C |
| ATOM | 1982 | O | GLY | B | 25 | 34.735 | 3.069 | -5.061 | 1.00 | 41.00 | B | O |
| ATOM | 1983 | N | LEU | B | 26 | 36.579 | 3.599 | -3.888 | 1.00 | 39.39 | B | N |
| ATOM | 1984 | CA | LEU | B | 26 | 36.656 | 2.302 | -3.237 | 1.00 | 38.36 | B | C |
| ATOM | 1985 | CB | LEU | B | 26 | 37.760 | 2.316 | -2.182 | 1.00 | 36.59 | B | C |
| ATOM | 1986 | CG | LEU | B | 26 | 39.174 | 2.447 | -2.745 | 1.00 | 34.96 | B | C |
| ATOM | 1987 | CD1 | LEU | B | 26 | 40.120 | 2.868 | -1.644 | 1.00 | 35.92 | B | C |
| ATOM | 1988 | CD2 | LEU | B | 26 | 39.604 | 1.135 | -3.365 | 1.00 | 32.64 | B | C |
| ATOM | 1989 | C | LEU | B | 26 | 35.338 | 1.945 | -2.578 | 1.00 | 39.11 | B | C |
| ATOM | 1990 | O | LEU | B | 26 | 34.891 | 0.804 | -2.642 | 1.00 | 38.61 | B | O |
| ATOM | 1991 | N | ARG | B | 27 | 34.715 | 2.932 | -1.944 | 1.00 | 40.50 | B | N |
| ATOM | 1992 | CA | ARG | B | 27 | 33.456 | 2.710 | -1.245 | 1.00 | 42.03 | B | C |
| ATOM | 1993 | CB | ARG | B | 27 | 33.084 | 3.943 | -0.427 | 1.00 | 42.52 | B | C |
| ATOM | 1994 | CG | ARG | B | 27 | 33.985 | 4.189 | 0.768 | 1.00 | 41.95 | B | C |
| ATOM | 1995 | CD | ARG | B | 27 | 33.460 | 5.370 | 1.543 | 1.00 | 43.49 | B | C |
| ATOM | 1996 | NE | ARG | B | 27 | 32.043 | 5.187 | 1.840 | 1.00 | 42.35 | B | N |
| ATOM | 1997 | CZ | ARG | B | 27 | 31.577 | 4.473 | 2.854 | 1.00 | 42.73 | B | C |
| ATOM | 1998 | NH1 | ARG | B | 27 | 32.415 | 3.869 | 3.692 | 1.00 | 43.63 | B | N |
| ATOM | 1999 | NH2 | ARG | B | 27 | 30.265 | 4.357 | 3.025 | 1.00 | 43.90 | B | N |
| ATOM | 2000 | C | ARG | B | 27 | 32.309 | 2.356 | -2.169 | 1.00 | 43.02 | B | C |
| ATOM | 2001 | O | ARG | B | 27 | 31.407 | 1.621 | -1.780 | 1.00 | 43.65 | B | O |
| ATOM | 2002 | N | GLN | B | 28 | 32.336 | 2.878 | -3.389 | 1.00 | 44.21 | B | N |
| ATOM | 2003 | CA | GLN | B | 28 | 31.281 | 2.583 | -4.344 | 1.00 | 46.08 | B | C |
| ATOM | 2004 | CB | GLN | B | 28 | 31.312 | 3.599 | -5.496 | 1.00 | 46.77 | B | C |
| ATOM | 2005 | CG | GLN | B | 28 | 30.352 | 4.796 | -5.298 | 1.00 | 49.00 | B | C |
| ATOM | 2006 | CD | GLN | B | 28 | 28.908 | 4.502 | -5.762 | 1.00 | 50.46 | B | C |
| ATOM | 2007 | OE1 | GLN | B | 28 | 27.951 | 5.151 | -5.317 | 1.00 | 49.49 | B | O |
| ATOM | 2008 | NE2 | GLN | B | 28 | 28.757 | 3.536 | -6.672 | 1.00 | 49.94 | B | N |
| ATOM | 2009 | C | GLN | B | 28 | 31.399 | 1.153 | -4.889 | 1.00 | 47.00 | B | C |
| ATOM | 2010 | O | GLN | B | 28 | 30.391 | 0.454 | -5.032 | 1.00 | 47.29 | B | O |
| ATOM | 2011 | N | LEU | B | 29 | 32.626 | 0.718 | -5.172 | 1.00 | 46.17 | B | N |
| ATOM | 2012 | CA | LEU | B | 29 | 32.875 | -0.622 | -5.718 | 1.00 | 45.38 | B | C |
| ATOM | 2013 | CB | LEU | B | 29 | 34.346 | -1.018 | -5.521 | 1.00 | 45.09 | B | C |
| ATOM | 2014 | CG | LEU | B | 29 | 35.444 | -0.186 | -6.171 | 1.00 | 45.32 | B | C |
| ATOM | 2015 | CD1 | LEU | B | 29 | 36.799 | -0.782 | -5.796 | 1.00 | 45.30 | B | C |
| ATOM | 2016 | CD2 | LEU | B | 29 | 35.280 | -0.176 | -7.681 | 1.00 | 46.01 | B | C |
| ATOM | 2017 | C | LEU | B | 29 | 32.023 | -1.742 | -5.137 | 1.00 | 44.63 | B | C |
| ATOM | 2018 | O | LEU | B | 29 | 31.690 | -1.740 | -3.952 | 1.00 | 43.46 | B | O |
| ATOM | 2019 | N | ALA | B | 30 | 31.675 | -2.708 | -5.984 | 1.00 | 44.88 | B | N |
| ATOM | 2020 | CA | ALA | B | 30 | 30.909 | -3.873 | -5.541 | 1.00 | 45.37 | B | C |
| ATOM | 2021 | CB | ALA | B | 30 | 29.766 | -4.174 | -6.518 | 1.00 | 45.11 | B | C |
| ATOM | 2022 | C | ALA | B | 30 | 31.938 | -5.004 | -5.552 | 1.00 | 45.01 | B | C |
| ATOM | 2023 | O | ALA | B | 30 | 32.496 | -5.314 | -6.603 | 1.00 | 45.84 | B | O |
| ATOM | 2024 | N | LEU | B | 31 | 32.196 | -5.606 | -4.393 | 1.00 | 44.06 | B | N |
| ATOM | 2025 | CA | LEU | B | 31 | 33.193 | -6.665 | -4.292 | 1.00 | 44.32 | B | C |
| ATOM | 2026 | CB | LEU | B | 31 | 34.095 | -6.376 | -3.087 | 1.00 | 44.80 | B | C |
| ATOM | 2027 | CG | LEU | B | 31 | 34.488 | -4.892 | -2.933 | 1.00 | 45.35 | B | C |
| ATOM | 2028 | CD1 | LEU | B | 31 | 35.087 | -4.654 | -1.545 | 1.00 | 44.66 | B | C |
| ATOM | 2029 | CD2 | LEU | B | 31 | 35.466 | -4.483 | -4.026 | 1.00 | 43.74 | B | C |
| ATOM | 2030 | C | LEU | B | 31 | 32.588 | -8.065 | -4.172 | 1.00 | 44.81 | B | C |
| ATOM | 2031 | O | LEU | B | 31 | 31.431 | -8.223 | -3.772 | 1.00 | 45.38 | B | O |
| ATOM | 2032 | N | PRO | B | 32 | 33.369 | -9.104 | -4.515 | 1.00 | 44.52 | B | N |
| ATOM | 2033 | CD | PRO | B | 32 | 34.652 | -9.011 | -5.229 | 1.00 | 44.20 | B | C |
| ATOM | 2034 | CA | PRO | B | 32 | 32.921 | -10.503 | -4.452 | 1.00 | 44.93 | B | C |
| ATOM | 2035 | CB | PRO | B | 32 | 34.034 | -11.250 | -5.195 | 1.00 | 45.12 | B | C |
| ATOM | 2036 | CG | PRO | B | 32 | 34.581 | -10.211 | -6.128 | 1.00 | 42.82 | B | C |
| ATOM | 2037 | C | PRO | B | 32 | 32.739 | -11.021 | -3.020 | 1.00 | 45.69 | B | C |
| ATOM | 2038 | O | PRO | B | 32 | 33.475 | -10.638 | -2.113 | 1.00 | 45.77 | B | O |
| ATOM | 2039 | N | SER | B | 33 | 31.765 | -11.902 | -2.816 | 1.00 | 45.92 | B | N |
| ATOM | 2040 | CA | SER | B | 33 | 31.512 | -12.437 | -1.480 | 1.00 | 46.21 | B | C |
| ATOM | 2041 | CB | SER | B | 33 | 30.258 | -13.321 | -1.481 | 1.00 | 47.75 | B | C |
| ATOM | 2042 | OG | SER | B | 33 | 30.111 | -13.961 | -0.216 | 1.00 | 48.32 | B | O |
| ATOM | 2043 | C | SER | B | 33 | 32.659 | -13.253 | -0.890 | 1.00 | 45.04 | B | C |

FIGURE 9 - 31

41/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|----|--------|---------|--------|------|-------|---|---|
| ATOM | 2044 | O | SER | B | 33 | 32.939 | -13.166 | 0.319 | 1.00 | 44.91 | B | O |
| ATOM | 2045 | N | GLY | B | 34 | 33.314 | -14.053 | -1.726 | 1.00 | 43.46 | B | N |
| ATOM | 2046 | CA | GLY | B | 34 | 34.398 | -14.884 | -1.219 | 1.00 | 43.32 | B | C |
| ATOM | 2047 | C | GLY | B | 34 | 35.809 | -14.348 | -1.370 | 1.00 | 41.91 | B | C |
| ATOM | 2048 | O | GLY | B | 34 | 36.766 | -15.129 | -1.424 | 1.00 | 42.84 | B | O |
| ATOM | 2049 | N | VAL | B | 35 | 35.956 | -13.027 | -1.408 | 1.00 | 40.61 | B | N |
| ATOM | 2050 | CA | VAL | B | 35 | 37.268 | -12.415 | -1.583 | 1.00 | 38.68 | B | C |
| ATOM | 2051 | CB | VAL | B | 35 | 37.380 | -11.799 | -2.999 | 1.00 | 39.03 | B | C |
| ATOM | 2052 | CG1 | VAL | B | 35 | 38.724 | -11.128 | -3.193 | 1.00 | 38.80 | B | C |
| ATOM | 2053 | CG2 | VAL | B | 35 | 37.191 | -12.880 | -4.037 | 1.00 | 40.23 | B | C |
| ATOM | 2054 | C | VAL | B | 35 | 37.627 | -11.340 | -0.560 | 1.00 | 37.38 | B | C |
| ATOM | 2055 | O | VAL | B | 35 | 36.962 | -10.311 | -0.469 | 1.00 | 37.36 | B | O |
| ATOM | 2056 | N | ASP | B | 36 | 38.679 | -11.590 | 0.211 | 1.00 | 35.95 | B | N |
| ATOM | 2057 | CA | ASP | B | 36 | 39.163 | -10.626 | 1.192 | 1.00 | 35.03 | B | C |
| ATOM | 2058 | CB | ASP | B | 36 | 40.201 | -11.275 | 2.112 | 1.00 | 36.19 | B | C |
| ATOM | 2059 | CG | ASP | B | 36 | 39.594 | -12.253 | 3.101 | 1.00 | 37.41 | B | C |
| ATOM | 2060 | OD1 | ASP | B | 36 | 40.337 | -13.138 | 3.577 | 1.00 | 38.15 | B | O |
| ATOM | 2061 | OD2 | ASP | B | 36 | 38.391 | -12.139 | 3.422 | 1.00 | 38.33 | B | O |
| ATOM | 2062 | C | ASP | B | 36 | 39.846 | -9.508 | 0.394 | 1.00 | 34.14 | B | C |
| ATOM | 2063 | O | ASP | B | 36 | 40.671 | -9.776 | -0.478 | 1.00 | 33.18 | B | O |
| ATOM | 2064 | N | VAL | B | 37 | 39.499 | -8.260 | 0.686 | 1.00 | 33.20 | B | N |
| ATOM | 2065 | CA | VAL | B | 37 | 40.087 | -7.119 | -0.009 | 1.00 | 31.64 | B | C |
| ATOM | 2066 | CB | VAL | B | 37 | 39.013 | -6.368 | -0.837 | 1.00 | 31.97 | B | C |
| ATOM | 2067 | CG1 | VAL | B | 37 | 39.600 | -5.124 | -1.474 | 1.00 | 31.19 | B | C |
| ATOM | 2068 | CG2 | VAL | B | 37 | 38.457 | -7.289 | -1.913 | 1.00 | 31.37 | B | C |
| ATOM | 2069 | C | VAL | B | 37 | 40.724 | -6.178 | 1.018 | 1.00 | 32.37 | B | C |
| ATOM | 2070 | O | VAL | B | 37 | 40.113 | -5.838 | 2.033 | 1.00 | 31.75 | B | O |
| ATOM | 2071 | N | ALA | B | 38 | 41.961 | -5.766 | 0.750 | 1.00 | 31.13 | B | N |
| ATOM | 2072 | CA | ALA | B | 38 | 42.687 | -4.884 | 1.657 | 1.00 | 29.80 | B | C |
| ATOM | 2073 | CB | ALA | B | 38 | 43.627 | -5.717 | 2.523 | 1.00 | 30.23 | B | C |
| ATOM | 2074 | C | ALA | B | 38 | 43.487 | -3.805 | 0.922 | 1.00 | 28.30 | B | C |
| ATOM | 2075 | O | ALA | B | 38 | 43.798 | -3.949 | -0.258 | 1.00 | 27.46 | B | O |
| ATOM | 2076 | N | VAL | B | 39 | 43.812 | -2.722 | 1.623 | 1.00 | 27.33 | B | N |
| ATOM | 2077 | CA | VAL | B | 39 | 44.615 | -1.648 | 1.036 | 1.00 | 27.10 | B | C |
| ATOM | 2078 | CB | VAL | B | 39 | 43.766 | -0.402 | 0.692 | 1.00 | 26.85 | B | C |
| ATOM | 2079 | CG1 | VAL | B | 39 | 42.813 | -0.730 | -0.429 | 1.00 | 26.50 | B | C |
| ATOM | 2080 | CG2 | VAL | B | 39 | 43.002 | 0.073 | 1.924 | 1.00 | 26.89 | B | C |
| ATOM | 2081 | C | VAL | B | 39 | 45.722 | -1.254 | 2.012 | 1.00 | 26.24 | B | C |
| ATOM | 2082 | O | VAL | B | 39 | 45.533 | -1.290 | 3.227 | 1.00 | 26.15 | B | O |
| ATOM | 2083 | N | MET | B | 40 | 46.880 | -0.895 | 1.470 | 1.00 | 25.21 | B | N |
| ATOM | 2084 | CA | MET | B | 40 | 48.034 | -0.504 | 2.280 | 1.00 | 24.34 | B | C |
| ATOM | 2085 | CB | MET | B | 40 | 49.247 | -1.369 | 1.923 | 1.00 | 23.60 | B | C |
| ATOM | 2086 | CG | MET | B | 40 | 49.313 | -2.739 | 2.562 | 1.00 | 26.47 | B | C |
| ATOM | 2087 | SD | MET | B | 40 | 48.033 | -3.883 | 2.079 | 1.00 | 25.47 | B | S |
| ATOM | 2088 | CE | MET | B | 40 | 47.533 | -4.394 | 3.603 | 1.00 | 22.65 | B | C |
| ATOM | 2089 | C | MET | B | 40 | 48.396 | 0.968 | 2.042 | 1.00 | 23.74 | B | C |
| ATOM | 2090 | O | MET | B | 40 | 49.330 | 1.276 | 1.294 | 1.00 | 22.64 | B | O |
| ATOM | 2091 | N | PRO | B | 41 | 47.663 | 1.900 | 2.672 | 1.00 | 23.57 | B | N |
| ATOM | 2092 | CD | PRO | B | 41 | 46.465 | 1.746 | 3.518 | 1.00 | 22.37 | B | C |
| ATOM | 2093 | CA | PRO | B | 41 | 47.986 | 3.316 | 2.466 | 1.00 | 22.75 | B | C |
| ATOM | 2094 | CB | PRO | B | 41 | 46.753 | 4.030 | 3.002 | 1.00 | 23.05 | B | C |
| ATOM | 2095 | CG | PRO | B | 41 | 46.291 | 3.135 | 4.083 | 1.00 | 23.95 | B | C |
| ATOM | 2096 | C | PRO | B | 41 | 49.248 | 3.727 | 3.207 | 1.00 | 22.63 | B | C |
| ATOM | 2097 | O | PRO | B | 41 | 49.689 | 3.040 | 4.127 | 1.00 | 19.56 | B | O |
| ATOM | 2098 | N | PRO | B | 42 | 49.860 | 4.848 | 2.798 | 1.00 | 23.12 | B | N |
| ATOM | 2099 | CD | PRO | B | 42 | 49.516 | 5.719 | 1.666 | 1.00 | 22.76 | B | C |
| ATOM | 2100 | CA | PRO | B | 42 | 51.077 | 5.320 | 3.464 | 1.00 | 23.38 | B | C |
| ATOM | 2101 | CB | PRO | B | 42 | 51.463 | 6.559 | 2.661 | 1.00 | 23.45 | B | C |
| ATOM | 2102 | CG | PRO | B | 42 | 50.853 | 6.310 | 1.322 | 1.00 | 23.94 | B | C |
| ATOM | 2103 | C | PRO | B | 42 | 50.665 | 5.682 | 4.878 | 1.00 | 22.81 | B | C |
| ATOM | 2104 | O | PRO | B | 42 | 49.555 | 6.154 | 5.100 | 1.00 | 23.49 | B | O |
| ATOM | 2105 | N | CYS | B | 43 | 51.556 | 5.450 | 5.824 | 1.00 | 23.09 | B | N |
| ATOM | 2106 | CA | CYS | B | 43 | 51.295 | 5.737 | 7.222 | 1.00 | 25.85 | B | C |
| ATOM | 2107 | CB | CYS | B | 43 | 52.627 | 5.745 | 7.974 | 1.00 | 29.28 | B | C |
| ATOM | 2108 | SG | CYS | B | 43 | 52.471 | 5.579 | 9.763 | 1.00 | 38.46 | B | S |
| ATOM | 2109 | C | CYS | B | 43 | 50.530 | 7.061 | 7.451 | 1.00 | 25.80 | B | C |

FIGURE 9 - 32

SUBSTITUTE SHEET (RULE 26)

42/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 2110 | O | CYS | B | 43 | 49.560 | 7.098 | 8.207 | 1.00 | 25.78 | B | O |
| ATOM | 2111 | N | LEU | B | 44 | 50.968 | 8.129 | 6.789 | 1.00 | 25.46 | B | N |
| ATOM | 2112 | CA | LEU | B | 44 | 50.348 | 9.453 | 6.903 | 1.00 | 25.93 | B | C |
| ATOM | 2113 | CB | LEU | B | 44 | 50.945 | 10.427 | 5.882 | 1.00 | 26.44 | B | C |
| ATOM | 2114 | CG | LEU | B | 44 | 51.937 | 11.515 | 6.276 | 1.00 | 27.42 | B | C |
| ATOM | 2115 | CD1 | LEU | B | 44 | 52.042 | 12.490 | 5.105 | 1.00 | 26.57 | B | C |
| ATOM | 2116 | CD2 | LEU | B | 44 | 51.475 | 12.250 | 7.527 | 1.00 | 26.20 | B | C |
| ATOM | 2117 | C | LEU | B | 44 | 48.846 | 9.465 | 6.682 | 1.00 | 25.07 | B | C |
| ATOM | 2118 | O | LEU | B | 44 | 48.160 | 10.333 | 7.198 | 1.00 | 24.58 | B | O |
| ATOM | 2119 | N | PHE | B | 45 | 48.337 | 8.514 | 5.909 | 1.00 | 24.12 | B | N |
| ATOM | 2120 | CA | PHE | B | 45 | 46.919 | 8.497 | 5.608 | 1.00 | 23.86 | B | C |
| ATOM | 2121 | CB | PHE | B | 45 | 46.728 | 8.626 | 4.093 | 1.00 | 23.72 | B | C |
| ATOM | 2122 | CG | PHE | B | 45 | 47.506 | 9.756 | 3.483 | 1.00 | 23.20 | B | C |
| ATOM | 2123 | CD1 | PHE | B | 45 | 48.714 | 9.514 | 2.836 | 1.00 | 21.67 | B | C |
| ATOM | 2124 | CD2 | PHE | B | 45 | 47.049 | 11.066 | 3.585 | 1.00 | 21.85 | B | C |
| ATOM | 2125 | CE1 | PHE | B | 45 | 49.454 | 10.555 | 2.303 | 1.00 | 20.18 | B | C |
| ATOM | 2126 | CE2 | PHE | B | 45 | 47.780 | 12.112 | 3.055 | 1.00 | 22.29 | B | C |
| ATOM | 2127 | CZ | PHE | B | 45 | 48.988 | 11.860 | 2.411 | 1.00 | 22.16 | B | C |
| ATOM | 2128 | C | PHE | B | 45 | 46.128 | 7.301 | 6.121 | 1.00 | 23.33 | B | C |
| ATOM | 2129 | O | PHE | B | 45 | 44.947 | 7.171 | 5.809 | 1.00 | 23.55 | B | O |
| ATOM | 2130 | N | ILE | B | 46 | 46.765 | 6.447 | 6.915 | 1.00 | 22.52 | B | N |
| ATOM | 2131 | CA | ILE | B | 46 | 46.098 | 5.265 | 7.439 | 1.00 | 22.50 | B | C |
| ATOM | 2132 | CB | ILE | B | 46 | 47.058 | 4.433 | 8.321 | 1.00 | 20.37 | B | C |
| ATOM | 2133 | CG2 | ILE | B | 46 | 46.275 | 3.517 | 9.224 | 1.00 | 20.98 | B | C |
| ATOM | 2134 | CG1 | ILE | B | 46 | 48.016 | 3.632 | 7.427 | 1.00 | 22.47 | B | C |
| ATOM | 2135 | CD1 | ILE | B | 46 | 48.986 | 2.731 | 8.192 | 1.00 | 20.92 | B | C |
| ATOM | 2136 | C | ILE | B | 46 | 44.799 | 5.537 | 8.212 | 1.00 | 23.39 | B | C |
| ATOM | 2137 | O | ILE | B | 46 | 43.771 | 4.920 | 7.918 | 1.00 | 20.35 | B | O |
| ATOM | 2138 | N | SER | B | 47 | 44.827 | 6.446 | 9.188 | 1.00 | 23.45 | B | N |
| ATOM | 2139 | CA | SER | B | 47 | 43.608 | 6.714 | 9.951 | 1.00 | 25.50 | B | C |
| ATOM | 2140 | CB | SER | B | 47 | 43.912 | 7.552 | 11.200 | 1.00 | 24.79 | B | C |
| ATOM | 2141 | OG | SER | B | 47 | 44.629 | 8.717 | 10.864 | 1.00 | 31.00 | B | O |
| ATOM | 2142 | C | SER | B | 47 | 42.540 | 7.385 | 9.080 | 1.00 | 25.78 | B | C |
| ATOM | 2143 | O | SER | B | 47 | 41.346 | 7.165 | 9.275 | 1.00 | 26.07 | B | O |
| ATOM | 2144 | N | GLN | B | 48 | 42.972 | 8.177 | 8.104 | 1.00 | 26.97 | B | N |
| ATOM | 2145 | CA | GLN | B | 48 | 42.045 | 8.838 | 7.196 | 1.00 | 27.38 | B | C |
| ATOM | 2146 | CB | GLN | B | 48 | 42.803 | 9.717 | 6.205 | 1.00 | 27.45 | B | C |
| ATOM | 2147 | CG | GLN | B | 48 | 41.904 | 10.585 | 5.321 | 1.00 | 30.13 | B | C |
| ATOM | 2148 | CD | GLN | B | 48 | 42.656 | 11.257 | 4.188 | 1.00 | 28.86 | B | C |
| ATOM | 2149 | OE1 | GLN | B | 48 | 43.831 | 11.605 | 4.326 | 1.00 | 31.46 | B | O |
| ATOM | 2150 | NE2 | GLN | B | 48 | 41.978 | 11.456 | 3.064 | 1.00 | 29.55 | B | N |
| ATOM | 2151 | C | GLN | B | 48 | 41.298 | 7.751 | 6.418 | 1.00 | 29.31 | B | C |
| ATOM | 2152 | O | GLN | B | 48 | 40.064 | 7.772 | 6.305 | 1.00 | 29.03 | B | O |
| ATOM | 2153 | N | VAL | B | 49 | 42.061 | 6.801 | 5.882 | 1.00 | 28.34 | B | N |
| ATOM | 2154 | CA | VAL | B | 49 | 41.489 | 5.711 | 5.100 | 1.00 | 27.85 | B | C |
| ATOM | 2155 | CB | VAL | B | 49 | 42.604 | 4.823 | 4.486 | 1.00 | 27.42 | B | C |
| ATOM | 2156 | CG1 | VAL | B | 49 | 42.043 | 3.468 | 4.064 | 1.00 | 25.19 | B | C |
| ATOM | 2157 | CG2 | VAL | B | 49 | 43.197 | 5.529 | 3.266 | 1.00 | 26.55 | B | C |
| ATOM | 2158 | C | VAL | B | 49 | 40.544 | 4.864 | 5.934 | 1.00 | 28.64 | B | C |
| ATOM | 2159 | O | VAL | B | 49 | 39.446 | 4.537 | 5.495 | 1.00 | 28.16 | B | O |
| ATOM | 2160 | N | ILE | B | 50 | 40.978 | 4.510 | 7.138 | 1.00 | 29.96 | B | N |
| ATOM | 2161 | CA | ILE | B | 50 | 40.156 | 3.715 | 8.035 | 1.00 | 31.54 | B | C |
| ATOM | 2162 | CB | ILE | B | 50 | 40.939 | 3.385 | 9.313 | 1.00 | 31.44 | B | C |
| ATOM | 2163 | CG2 | ILE | B | 50 | 39.995 | 2.936 | 10.423 | 1.00 | 30.95 | B | C |
| ATOM | 2164 | CG1 | ILE | B | 50 | 41.969 | 2.302 | 8.987 | 1.00 | 31.14 | B | C |
| ATOM | 2165 | CD1 | ILE | B | 50 | 42.959 | 2.043 | 10.083 | 1.00 | 33.49 | B | C |
| ATOM | 2166 | C | ILE | B | 50 | 38.853 | 4.443 | 8.370 | 1.00 | 33.33 | B | C |
| ATOM | 2167 | O | ILE | B | 50 | 37.763 | 3.884 | 8.209 | 1.00 | 33.92 | B | O |
| ATOM | 2168 | N | GLN | B | 51 | 38.950 | 5.684 | 8.839 | 1.00 | 34.64 | B | N |
| ATOM | 2169 | CA | GLN | B | 51 | 37.738 | 6.433 | 9.137 | 1.00 | 36.40 | B | C |
| ATOM | 2170 | CB | GLN | B | 51 | 38.045 | 7.902 | 9.458 | 1.00 | 38.21 | B | C |
| ATOM | 2171 | CG | GLN | B | 51 | 38.580 | 8.166 | 10.862 | 1.00 | 43.10 | B | C |
| ATOM | 2172 | CD | GLN | B | 51 | 38.751 | 9.663 | 11.163 | 1.00 | 46.44 | B | C |
| ATOM | 2173 | OE1 | GLN | B | 51 | 39.322 | 10.041 | 12.201 | 1.00 | 48.54 | B | O |
| ATOM | 2174 | NE2 | GLN | B | 51 | 38.253 | 10.518 | 10.263 | 1.00 | 47.05 | B | N |
| ATOM | 2175 | C | GLN | B | 51 | 36.888 | 6.384 | 7.873 | 1.00 | 35.62 | B | C |

FIGURE 9- 33

SUBSTITUTE SHEET (RULE 26)

43/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 2176 | O | GLN | B | 51 | 35.723 | 6.019 | 7.915 | 1.00 | 36.83 | | |
| ATOM | 2177 | N | GLY | B | 52 | 37.505 | 6.734 | 6.747 | 1.00 | 35.41 | B | O |
| ATOM | 2178 | CA | GLY | B | 52 | 36.812 | 6.770 | 5.469 | 1.00 | 34.30 | B | N |
| ATOM | 2179 | C | GLY | B | 52 | 36.124 | 5.506 | 4.977 | 1.00 | 34.21 | B | C |
| ATOM | 2180 | O | GLY | B | 52 | 35.050 | 5.583 | 4.374 | 1.00 | 35.08 | B | O |
| ATOM | 2181 | N | LEU | B | 53 | 36.719 | 4.345 | 5.219 | 1.00 | 34.15 | B | N |
| ATOM | 2182 | CA | LEU | B | 53 | 36.125 | 3.096 | 4.758 | 1.00 | 33.93 | B | C |
| ATOM | 2183 | CB | LEU | B | 53 | 37.215 | 2.161 | 4.219 | 1.00 | 32.97 | B | C |
| ATOM | 2184 | CG | LEU | B | 53 | 37.948 | 2.620 | 2.962 | 1.00 | 31.86 | B | C |
| ATOM | 2185 | CD1 | LEU | B | 53 | 38.983 | 1.570 | 2.571 | 1.00 | 31.41 | B | C |
| ATOM | 2186 | CD2 | LEU | B | 53 | 36.953 | 2.844 | 1.833 | 1.00 | 31.36 | B | C |
| ATOM | 2187 | C | LEU | B | 53 | 35.298 | 2.351 | 5.802 | 1.00 | 35.02 | B | C |
| ATOM | 2188 | O | LEU | B | 53 | 34.961 | 1.182 | 5.600 | 1.00 | 35.57 | B | O |
| ATOM | 2189 | N | ALA | B | 54 | 34.971 | 3.002 | 6.916 | 1.00 | 34.86 | B | N |
| ATOM | 2190 | CA | ALA | B | 54 | 34.164 | 2.343 | 7.940 | 1.00 | 35.40 | B | C |
| ATOM | 2191 | CB | ALA | B | 54 | 33.894 | 3.304 | 9.107 | 1.00 | 36.74 | B | C |
| ATOM | 2192 | C | ALA | B | 54 | 32.842 | 1.892 | 7.300 | 1.00 | 35.40 | B | C |
| ATOM | 2193 | O | ALA | B | 54 | 32.267 | 2.602 | 6.470 | 1.00 | 33.69 | B | O |
| ATOM | 2194 | N | GLY | B | 55 | 32.379 | 0.702 | 7.671 | 1.00 | 36.38 | B | N |
| ATOM | 2195 | CA | GLY | B | 55 | 31.145 | 0.187 | 7.106 | 1.00 | 37.79 | B | C |
| ATOM | 2196 | C | GLY | B | 55 | 31.353 | -0.695 | 5.881 | 1.00 | 39.23 | B | C |
| ATOM | 2197 | O | GLY | B | 55 | 30.442 | -1.419 | 5.474 | 1.00 | 39.28 | B | O |
| ATOM | 2198 | N | LYS | B | 56 | 32.542 | -0.635 | 5.285 | 1.00 | 39.48 | B | N |
| ATOM | 2199 | CA | LYS | B | 56 | 32.850 | -1.444 | 4.108 | 1.00 | 39.66 | B | C |
| ATOM | 2200 | CB | LYS | B | 56 | 33.561 | -0.607 | 3.044 | 1.00 | 40.26 | B | C |
| ATOM | 2201 | CG | LYS | B | 56 | 32.662 | 0.358 | 2.307 | 1.00 | 42.60 | B | C |
| ATOM | 2202 | CD | LYS | B | 56 | 31.589 | -0.366 | 1.479 | 1.00 | 44.60 | B | C |
| ATOM | 2203 | CE | LYS | B | 56 | 32.203 | -1.248 | 0.392 | 1.00 | 45.54 | B | C |
| ATOM | 2204 | NZ | LYS | B | 56 | 33.130 | -0.480 | -0.510 | 1.00 | 45.54 | B | N |
| ATOM | 2205 | C | LYS | B | 56 | 33.726 | -2.634 | 4.474 | 1.00 | 39.17 | B | C |
| ATOM | 2206 | O | LYS | B | 56 | 34.628 | -2.522 | 5.301 | 1.00 | 39.65 | B | O |
| ATOM | 2207 | N | ALA | B | 57 | 33.464 | -3.776 | 3.852 | 1.00 | 38.57 | B | N |
| ATOM | 2208 | CA | ALA | B | 57 | 34.243 | -4.970 | 4.124 | 1.00 | 37.40 | B | C |
| ATOM | 2209 | CB | ALA | B | 57 | 33.455 | -6.207 | 3.715 | 1.00 | 37.30 | B | C |
| ATOM | 2210 | C | ALA | B | 57 | 35.565 | -4.895 | 3.364 | 1.00 | 37.00 | B | C |
| ATOM | 2211 | O | ALA | B | 57 | 35.833 | -5.705 | 2.475 | 1.00 | 38.04 | B | O |
| ATOM | 2212 | N | ILE | B | 58 | 36.369 | -3.890 | 3.705 | 1.00 | 35.62 | B | N |
| ATOM | 2213 | CA | ILE | B | 58 | 37.688 | -3.678 | 3.105 | 1.00 | 34.47 | B | C |
| ATOM | 2214 | CB | ILE | B | 58 | 37.723 | -2.430 | 2.207 | 1.00 | 34.20 | B | C |
| ATOM | 2215 | CG2 | ILE | B | 58 | 39.169 | -2.054 | 1.900 | 1.00 | 33.74 | B | C |
| ATOM | 2216 | CG1 | ILE | B | 58 | 36.945 | -2.688 | 0.917 | 1.00 | 33.60 | B | C |
| ATOM | 2217 | CD1 | ILE | B | 58 | 36.881 | -1.489 | -0.006 | 1.00 | 32.56 | B | C |
| ATOM | 2218 | C | ILE | B | 58 | 38.681 | -3.458 | 4.241 | 1.00 | 34.23 | B | C |
| ATOM | 2219 | O | ILE | B | 58 | 38.541 | -2.512 | 5.015 | 1.00 | 34.54 | B | O |
| ATOM | 2220 | N | ASP | B | 59 | 39.673 | -4.331 | 4.355 | 1.00 | 33.60 | B | N |
| ATOM | 2221 | CA | ASP | B | 59 | 40.662 | -4.197 | 5.419 | 1.00 | 33.57 | B | C |
| ATOM | 2222 | CB | ASP | B | 59 | 41.273 | -5.560 | 5.759 | 1.00 | 34.73 | B | C |
| ATOM | 2223 | CG | ASP | B | 59 | 40.329 | -6.432 | 6.553 | 1.00 | 38.32 | B | C |
| ATOM | 2224 | OD1 | ASP | B | 59 | 39.685 | -5.910 | 7.491 | 1.00 | 39.68 | B | O |
| ATOM | 2225 | OD2 | ASP | B | 59 | 40.234 | -7.641 | 6.254 | 1.00 | 40.20 | B | O |
| ATOM | 2226 | C | ASP | B | 59 | 41.771 | -3.217 | 5.061 | 1.00 | 32.24 | B | C |
| ATOM | 2227 | O | ASP | B | 59 | 42.138 | -3.074 | 3.900 | 1.00 | 31.91 | B | O |
| ATOM | 2228 | N | VAL | B | 60 | 42.300 | -2.543 | 6.075 | 1.00 | 30.12 | B | N |
| ATOM | 2229 | CA | VAL | B | 60 | 43.374 | -1.579 | 5.884 | 1.00 | 27.95 | B | C |
| ATOM | 2230 | CB | VAL | B | 60 | 43.024 | -0.228 | 6.526 | 1.00 | 26.58 | B | C |
| ATOM | 2231 | CG1 | VAL | B | 60 | 44.227 | 0.699 | 6.464 | 1.00 | 26.55 | B | C |
| ATOM | 2232 | CG2 | VAL | B | 60 | 41.831 | 0.383 | 5.822 | 1.00 | 26.18 | B | C |
| ATOM | 2233 | C | VAL | B | 60 | 44.635 | -2.105 | 6.554 | 1.00 | 26.94 | B | C |
| ATOM | 2234 | O | VAL | B | 60 | 44.603 | -2.497 | 7.719 | 1.00 | 25.63 | B | O |
| ATOM | 2235 | N | GLY | B | 61 | 45.739 | -2.128 | 5.813 | 1.00 | 25.42 | B | N |
| ATOM | 2236 | CA | GLY | B | 61 | 46.990 | -2.596 | 6.380 | 1.00 | 26.08 | B | C |
| ATOM | 2237 | C | GLY | B | 61 | 48.074 | -1.541 | 6.207 | 1.00 | 25.45 | B | C |
| ATOM | 2238 | O | GLY | B | 61 | 47.820 | -0.491 | 5.634 | 1.00 | 25.17 | B | O |
| ATOM | 2239 | N | ALA | B | 62 | 49.278 | -1.816 | 6.699 | 1.00 | 25.32 | B | N |
| ATOM | 2240 | CA | ALA | B | 62 | 50.392 | -0.881 | 6.581 | 1.00 | 24.89 | B | C |
| ATOM | 2241 | CB | ALA | B | 62 | 50.979 | -0.601 | 7.959 | 1.00 | 24.37 | B | C |

FIGURE 9 - 34

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | | | | |
|------|------|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 2242 | C | ALA | B | 62 | 51.473 | -1.429 | 5.644 | 1.00 | 24.55 | B | C |
| ATOM | 2243 | O | ALA | B | 62 | 51.545 | -2.633 | 5.395 | 1.00 | 25.32 | B | O |
| ATOM | 2244 | N | GLN | B | 63 | 52.316 | -0.535 | 5.136 | 1.00 | 25.65 | B | N |
| ATOM | 2245 | CA | GLN | B | 63 | 53.388 | -0.905 | 4.213 | 1.00 | 23.99 | B | C |
| ATOM | 2246 | CB | GLN | B | 63 | 53.651 | 0.244 | 3.238 | 1.00 | 23.86 | B | C |
| ATOM | 2247 | CG | GLN | B | 63 | 52.510 | 0.551 | 2.285 | 1.00 | 22.80 | B | C |
| ATOM | 2248 | CD | GLN | B | 63 | 52.741 | 1.826 | 1.487 | 1.00 | 22.27 | B | C |
| ATOM | 2249 | OE1 | GLN | B | 63 | 53.826 | 2.050 | 0.953 | 1.00 | 25.18 | B | O |
| ATOM | 2250 | NE2 | GLN | B | 63 | 51.713 | 2.661 | 1.395 | 1.00 | 23.96 | B | N |
| ATOM | 2251 | C | GLN | B | 63 | 54.692 | -1.251 | 4.926 | 1.00 | 23.58 | B | C |
| ATOM | 2252 | O | GLN | B | 63 | 55.604 | -1.809 | 4.320 | 1.00 | 23.14 | B | O |
| ATOM | 2253 | N | ASN | B | 64 | 54.771 | -0.923 | 6.211 | 1.00 | 23.31 | B | N |
| ATOM | 2254 | CA | ASN | B | 64 | 55.975 | -1.172 | 7.007 | 1.00 | 23.10 | B | C |
| ATOM | 2255 | CB | ASN | B | 64 | 57.072 | -0.170 | 6.599 | 1.00 | 23.49 | B | C |
| ATOM | 2256 | CG | ASN | B | 64 | 58.389 | -0.390 | 7.341 | 1.00 | 23.95 | B | C |
| ATOM | 2257 | OD1 | ASN | B | 64 | 58.870 | -1.522 | 7.468 | 1.00 | 22.43 | B | O |
| ATOM | 2258 | ND2 | ASN | B | 64 | 58.986 | 0.699 | 7.818 | 1.00 | 21.25 | B | N |
| ATOM | 2259 | C | ASN | B | 64 | 55.653 | -1.017 | 8.491 | 1.00 | 22.24 | B | C |
| ATOM | 2260 | O | ASN | B | 64 | 54.641 | -0.419 | 8.842 | 1.00 | 22.35 | B | O |
| ATOM | 2261 | N | SER | B | 65 | 56.507 | -1.580 | 9.347 | 1.00 | 22.21 | B | N |
| ATOM | 2262 | CA | SER | B | 65 | 56.362 | -1.505 | 10.805 | 1.00 | 22.61 | B | C |
| ATOM | 2263 | CB | SER | B | 65 | 55.306 | -2.500 | 11.307 | 1.00 | 23.71 | B | C |
| ATOM | 2264 | OG | SER | B | 65 | 55.520 | -3.804 | 10.803 | 1.00 | 24.97 | B | O |
| ATOM | 2265 | C | SER | B | 65 | 57.718 | -1.771 | 11.477 | 1.00 | 23.40 | B | C |
| ATOM | 2266 | O | SER | B | 65 | 58.641 | -2.280 | 10.837 | 1.00 | 23.31 | B | O |
| ATOM | 2267 | N | ALA | B | 66 | 57.821 | -1.440 | 12.765 | 1.00 | 22.39 | B | N |
| ATOM | 2268 | CA | ALA | B | 66 | 59.066 | -1.565 | 13.537 | 1.00 | 23.04 | B | C |
| ATOM | 2269 | CB | ALA | B | 66 | 58.957 | -0.713 | 14.796 | 1.00 | 20.94 | B | C |
| ATOM | 2270 | C | ALA | B | 66 | 59.589 | -2.950 | 13.926 | 1.00 | 23.25 | B | C |
| ATOM | 2271 | O | ALA | B | 66 | 58.841 | -3.922 | 13.994 | 1.00 | 25.29 | B | O |
| ATOM | 2272 | N | VAL | B | 67 | 60.891 | -3.022 | 14.192 | 1.00 | 22.54 | B | N |
| ATOM | 2273 | CA | VAL | B | 67 | 61.510 | -4.272 | 14.612 | 1.00 | 21.56 | B | C |
| ATOM | 2274 | CB | VAL | B | 67 | 63.055 | -4.200 | 14.535 | 1.00 | 21.37 | B | C |
| ATOM | 2275 | CG1 | VAL | B | 67 | 63.499 | -4.181 | 13.079 | 1.00 | 20.67 | B | C |
| ATOM | 2276 | CG2 | VAL | B | 67 | 63.568 | -2.960 | 15.263 | 1.00 | 18.94 | B | C |
| ATOM | 2277 | C | VAL | B | 67 | 61.092 | -4.590 | 16.053 | 1.00 | 23.27 | B | C |
| ATOM | 2278 | O | VAL | B | 67 | 61.314 | -5.695 | 16.543 | 1.00 | 23.68 | B | O |
| ATOM | 2279 | N | GLU | B | 68 | 60.498 | -3.613 | 16.732 | 1.00 | 23.20 | B | N |
| ATOM | 2280 | CA | GLU | B | 68 | 60.039 | -3.814 | 18.099 | 1.00 | 24.95 | B | C |
| ATOM | 2281 | CB | GLU | B | 68 | 60.656 | -2.765 | 19.030 | 1.00 | 27.14 | B | C |
| ATOM | 2282 | CG | GLU | B | 68 | 62.183 | -2.721 | 18.948 | 1.00 | 32.62 | B | C |
| ATOM | 2283 | CD | GLU | B | 68 | 62.825 | -1.931 | 20.080 | 1.00 | 36.35 | B | C |
| ATOM | 2284 | OE1 | GLU | B | 68 | 62.275 | -0.871 | 20.466 | 1.00 | 35.99 | B | O |
| ATOM | 2285 | OE2 | GLU | B | 68 | 63.888 | -2.370 | 20.583 | 1.00 | 39.89 | B | O |
| ATOM | 2286 | C | GLU | B | 68 | 58.508 | -3.757 | 18.113 | 1.00 | 24.08 | B | C |
| ATOM | 2287 | O | GLU | B | 68 | 57.896 | -2.968 | 17.401 | 1.00 | 23.63 | B | O |
| ATOM | 2288 | N | PRO | B | 69 | 57.874 | -4.606 | 18.930 | 1.00 | 23.34 | B | N |
| ATOM | 2289 | CD | PRO | B | 69 | 58.534 | -5.617 | 19.774 | 1.00 | 23.98 | B | C |
| ATOM | 2290 | CA | PRO | B | 69 | 56.413 | -4.692 | 19.050 | 1.00 | 23.51 | B | C |
| ATOM | 2291 | CB | PRO | B | 69 | 56.210 | -6.037 | 19.733 | 1.00 | 24.45 | B | C |
| ATOM | 2292 | CG | PRO | B | 69 | 57.398 | -6.112 | 20.645 | 1.00 | 24.76 | B | C |
| ATOM | 2293 | C | PRO | B | 69 | 55.622 | -3.587 | 19.742 | 1.00 | 22.55 | B | C |
| ATOM | 2294 | O | PRO | B | 69 | 54.514 | -3.256 | 19.295 | 1.00 | 20.65 | B | O |
| ATOM | 2295 | N | MET | B | 70 | 56.168 | -3.028 | 20.822 | 1.00 | 21.13 | B | N |
| ATOM | 2296 | CA | MET | B | 70 | 55.456 | -1.996 | 21.575 | 1.00 | 20.07 | B | C |
| ATOM | 2297 | CB | MET | B | 70 | 55.440 | -2.363 | 23.062 | 1.00 | 19.64 | B | C |
| ATOM | 2298 | CG | MET | B | 70 | 54.888 | -3.755 | 23.355 | 1.00 | 21.95 | B | C |
| ATOM | 2299 | SD | MET | B | 70 | 53.144 | -3.966 | 22.919 | 1.00 | 22.78 | B | S |
| ATOM | 2300 | CE | MET | B | 70 | 52.468 | -4.249 | 24.527 | 1.00 | 23.64 | B | C |
| ATOM | 2301 | C | MET | B | 70 | 56.021 | -0.588 | 21.412 | 1.00 | 20.07 | B | C |
| ATOM | 2302 | O | MET | B | 70 | 57.109 | -0.398 | 20.874 | 1.00 | 16.62 | B | O |
| ATOM | 2303 | N | GLN | B | 71 | 55.268 | 0.398 | 21.888 | 1.00 | 19.62 | B | N |
| ATOM | 2304 | CA | GLN | B | 71 | 55.702 | 1.785 | 21.815 | 1.00 | 21.78 | B | C |
| ATOM | 2305 | CB | GLN | B | 71 | 54.638 | 2.699 | 22.430 | 1.00 | 21.26 | B | C |
| ATOM | 2306 | CG | GLN | B | 71 | 53.318 | 2.696 | 21.675 | 1.00 | 21.88 | B | C |
| ATOM | 2307 | CD | GLN | B | 71 | 52.249 | 3.497 | 22.381 | 1.00 | 25.66 | B | C |

FIGURE 9- 35

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | | | | | |
|------|------|-----|-----|---|----|--------|--------|--------|------|-------|---|---|---|
| ATOM | 2308 | OE1 | GLN | B | 71 | 52.471 | 4.654 | 22.751 | 1.00 | 25.01 | | | |
| ATOM | 2309 | NE2 | GLN | B | 71 | 51.080 | 2.891 | 22.576 | 1.00 | 24.70 | B | O | |
| ATOM | 2310 | C | GLN | B | 71 | 57.029 | 1.928 | 22.560 | 1.00 | 21.12 | B | N | C |
| ATOM | 2311 | O | GLN | B | 71 | 57.272 | 1.216 | 23.532 | 1.00 | 20.54 | B | O | |
| ATOM | 2312 | N | GLY | B | 72 | 57.893 | 2.823 | 22.090 | 1.00 | 22.14 | B | N | |
| ATOM | 2313 | CA | GLY | B | 72 | 59.177 | 3.011 | 22.744 | 1.00 | 24.64 | B | C | |
| ATOM | 2314 | C | GLY | B | 72 | 60.110 | 4.042 | 22.125 | 1.00 | 24.68 | B | C | |
| ATOM | 2315 | O | GLY | B | 72 | 59.689 | 4.909 | 21.363 | 1.00 | 23.09 | B | O | |
| ATOM | 2316 | N | ALA | B | 73 | 61.393 | 3.927 | 22.453 | 1.00 | 25.63 | B | N | |
| ATOM | 2317 | CA | ALA | B | 73 | 62.407 | 4.856 | 21.971 | 1.00 | 26.60 | B | C | |
| ATOM | 2318 | CB | ALA | B | 73 | 63.506 | 4.989 | 23.022 | 1.00 | 25.29 | B | C | |
| ATOM | 2319 | C | ALA | B | 73 | 63.025 | 4.516 | 20.604 | 1.00 | 26.95 | B | C | |
| ATOM | 2320 | O | ALA | B | 73 | 64.211 | 4.177 | 20.514 | 1.00 | 28.37 | B | O | |
| ATOM | 2321 | N | LEU | B | 74 | 62.223 | 4.616 | 19.548 | 1.00 | 24.96 | B | N | |
| ATOM | 2322 | CA | LEU | B | 74 | 62.681 | 4.353 | 18.187 | 1.00 | 23.24 | B | C | |
| ATOM | 2323 | CB | LEU | B | 74 | 62.188 | 2.987 | 17.692 | 1.00 | 23.03 | B | C | |
| ATOM | 2324 | CG | LEU | B | 74 | 63.057 | 1.744 | 17.924 | 1.00 | 24.92 | B | C | |
| ATOM | 2325 | CD1 | LEU | B | 74 | 62.419 | 0.541 | 17.223 | 1.00 | 22.55 | B | C | |
| ATOM | 2326 | CD2 | LEU | B | 74 | 64.471 | 1.996 | 17.386 | 1.00 | 21.17 | B | C | |
| ATOM | 2327 | C | LEU | B | 74 | 62.171 | 5.441 | 17.247 | 1.00 | 23.15 | B | C | |
| ATOM | 2328 | O | LEU | B | 74 | 61.194 | 5.248 | 16.530 | 1.00 | 23.97 | B | O | |
| ATOM | 2329 | N | THR | B | 75 | 62.846 | 6.583 | 17.266 | 1.00 | 22.29 | B | N | |
| ATOM | 2330 | CA | THR | B | 75 | 62.505 | 7.734 | 16.437 | 1.00 | 20.47 | B | C | |
| ATOM | 2331 | CB | THR | B | 75 | 63.629 | 8.779 | 16.500 | 1.00 | 19.93 | B | C | |
| ATOM | 2332 | OG1 | THR | B | 75 | 63.818 | 9.183 | 17.861 | 1.00 | 19.20 | B | O | |
| ATOM | 2333 | CG2 | THR | B | 75 | 63.291 | 9.979 | 15.637 | 1.00 | 17.55 | B | C | |
| ATOM | 2334 | C | THR | B | 75 | 62.210 | 7.454 | 14.961 | 1.00 | 20.89 | B | C | |
| ATOM | 2335 | O | THR | B | 75 | 63.006 | 6.821 | 14.254 | 1.00 | 17.45 | B | O | |
| ATOM | 2336 | N | GLY | B | 76 | 61.063 | 7.965 | 14.510 | 1.00 | 20.32 | B | N | |
| ATOM | 2337 | CA | GLY | B | 76 | 60.634 | 7.810 | 13.128 | 1.00 | 21.03 | B | C | |
| ATOM | 2338 | C | GLY | B | 76 | 60.052 | 6.459 | 12.762 | 1.00 | 21.09 | B | C | |
| ATOM | 2339 | O | GLY | B | 76 | 59.652 | 6.243 | 11.623 | 1.00 | 20.83 | B | O | |
| ATOM | 2340 | N | GLU | B | 77 | 59.987 | 5.562 | 13.738 | 1.00 | 20.47 | B | N | |
| ATOM | 2341 | CA | GLU | B | 77 | 59.489 | 4.198 | 13.545 | 1.00 | 22.74 | B | C | |
| ATOM | 2342 | CB | GLU | B | 77 | 60.444 | 3.268 | 14.290 | 1.00 | 24.89 | B | C | |
| ATOM | 2343 | CG | GLU | B | 77 | 60.920 | 2.026 | 13.586 | 1.00 | 29.85 | B | C | |
| ATOM | 2344 | CD | GLU | B | 77 | 61.391 | 2.277 | 12.177 | 1.00 | 30.41 | B | C | |
| ATOM | 2345 | OE1 | GLU | B | 77 | 62.037 | 3.317 | 11.921 | 1.00 | 33.21 | B | O | |
| ATOM | 2346 | OE2 | GLU | B | 77 | 61.113 | 1.413 | 11.321 | 1.00 | 33.54 | B | O | |
| ATOM | 2347 | C | GLU | B | 77 | 58.050 | 4.030 | 14.077 | 1.00 | 21.84 | B | C | |
| ATOM | 2348 | O | GLU | B | 77 | 57.682 | 4.642 | 15.073 | 1.00 | 21.17 | B | O | |
| ATOM | 2349 | N | THR | B | 78 | 57.241 | 3.206 | 13.410 | 1.00 | 23.32 | B | N | |
| ATOM | 2350 | CA | THR | B | 78 | 55.857 | 2.960 | 13.837 | 1.00 | 22.87 | B | C | |
| ATOM | 2351 | CB | THR | B | 78 | 54.860 | 3.033 | 12.666 | 1.00 | 24.53 | B | C | |
| ATOM | 2352 | OG1 | THR | B | 78 | 54.961 | 4.301 | 12.016 | 1.00 | 28.42 | B | O | |
| ATOM | 2353 | CG2 | THR | B | 78 | 53.431 | 2.853 | 13.178 | 1.00 | 23.72 | B | C | |
| ATOM | 2354 | C | THR | B | 78 | 55.706 | 1.572 | 14.447 | 1.00 | 22.95 | B | C | |
| ATOM | 2355 | O | THR | B | 78 | 55.965 | 0.562 | 13.787 | 1.00 | 22.65 | B | O | |
| ATOM | 2356 | N | ALA | B | 79 | 55.256 | 1.511 | 15.693 | 1.00 | 22.05 | B | N | |
| ATOM | 2357 | CA | ALA | B | 79 | 55.085 | 0.218 | 16.355 | 1.00 | 22.83 | B | C | |
| ATOM | 2358 | CB | ALA | B | 79 | 55.034 | 0.402 | 17.868 | 1.00 | 22.32 | B | C | |
| ATOM | 2359 | C | ALA | B | 79 | 53.843 | -0.546 | 15.880 | 1.00 | 22.52 | B | C | |
| ATOM | 2360 | O | ALA | B | 79 | 52.759 | 0.023 | 15.734 | 1.00 | 20.58 | B | O | |
| ATOM | 2361 | N | PRO | B | 80 | 53.993 | -1.857 | 15.630 | 1.00 | 22.09 | B | N | |
| ATOM | 2362 | CD | PRO | B | 80 | 55.219 | -2.666 | 15.724 | 1.00 | 20.76 | B | C | |
| ATOM | 2363 | CA | PRO | B | 80 | 52.866 | -2.677 | 15.177 | 1.00 | 21.33 | B | C | |
| ATOM | 2364 | CB | PRO | B | 80 | 53.489 | -4.066 | 15.015 | 1.00 | 21.52 | B | C | |
| ATOM | 2365 | CG | PRO | B | 80 | 54.654 | -4.049 | 15.937 | 1.00 | 21.13 | B | C | |
| ATOM | 2366 | C | PRO | B | 80 | 51.667 | -2.652 | 16.135 | 1.00 | 21.80 | B | C | |
| ATOM | 2367 | O | PRO | B | 80 | 50.524 | -2.824 | 15.706 | 1.00 | 22.64 | B | O | |
| ATOM | 2368 | N | SER | B | 81 | 51.918 | -2.431 | 17.425 | 1.00 | 22.29 | B | N | |
| ATOM | 2369 | CA | SER | B | 81 | 50.821 | -2.363 | 18.383 | 1.00 | 23.19 | B | C | |
| ATOM | 2370 | CB | SER | B | 81 | 51.334 | -2.371 | 19.826 | 1.00 | 25.28 | B | C | |
| ATOM | 2371 | OG | SER | B | 81 | 52.140 | -1.239 | 20.104 | 1.00 | 29.67 | B | O | |
| ATOM | 2372 | C | SER | B | 81 | 50.026 | -1.093 | 18.120 | 1.00 | 23.50 | B | C | |
| ATOM | 2373 | O | SER | B | 81 | 48.818 | -1.064 | 18.335 | 1.00 | 23.58 | B | O | |

FIGURE 9- 36

SUBSTITUTE SHEET (RULE 26)

46/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|----|--------|--------|--------|------|-------|---|---|
| ATOM | 2374 | N | GLN | B | 82 | 50.703 | -0.042 | 17.656 | 1.00 | 22.81 | B | N |
| ATOM | 2375 | CA | GLN | B | 82 | 50.023 | 1.213 | 17.344 | 1.00 | 23.53 | B | C |
| ATOM | 2376 | CB | GLN | B | 82 | 51.029 | 2.348 | 17.064 | 1.00 | 22.55 | B | C |
| ATOM | 2377 | CG | GLN | B | 82 | 51.823 | 2.817 | 18.283 | 1.00 | 21.14 | B | C |
| ATOM | 2378 | CD | GLN | B | 82 | 52.859 | 3.888 | 17.951 | 1.00 | 23.88 | B | C |
| ATOM | 2379 | OE1 | GLN | B | 82 | 53.787 | 3.651 | 17.179 | 1.00 | 24.02 | B | O |
| ATOM | 2380 | NE2 | GLN | B | 82 | 52.703 | 5.074 | 18.542 | 1.00 | 20.77 | B | N |
| ATOM | 2381 | C | GLN | B | 82 | 49.156 | 0.973 | 16.110 | 1.00 | 23.36 | B | C |
| ATOM | 2382 | O | GLN | B | 82 | 48.023 | 1.433 | 16.044 | 1.00 | 25.01 | B | O |
| ATOM | 2383 | N | LEU | B | 83 | 49.690 | 0.242 | 15.136 | 1.00 | 24.28 | B | N |
| ATOM | 2384 | CA | LEU | B | 83 | 48.948 | -0.060 | 13.916 | 1.00 | 23.92 | B | C |
| ATOM | 2385 | CB | LEU | B | 83 | 49.815 | -0.854 | 12.938 | 1.00 | 24.56 | B | C |
| ATOM | 2386 | CG | LEU | B | 83 | 50.886 | -0.056 | 12.194 | 1.00 | 25.89 | B | C |
| ATOM | 2387 | CD1 | LEU | B | 83 | 51.635 | -0.990 | 11.264 | 1.00 | 26.43 | B | C |
| ATOM | 2388 | CD2 | LEU | B | 83 | 50.238 | 1.095 | 11.413 | 1.00 | 24.09 | B | C |
| ATOM | 2389 | C | LEU | B | 83 | 47.700 | -0.854 | 14.257 | 1.00 | 24.04 | B | C |
| ATOM | 2390 | O | LEU | B | 83 | 46.626 | -0.586 | 13.740 | 1.00 | 25.72 | B | O |
| ATOM | 2391 | N | ALA | B | 84 | 47.850 | -1.836 | 15.135 | 1.00 | 23.98 | B | N |
| ATOM | 2392 | CA | ALA | B | 84 | 46.723 | -2.649 | 15.561 | 1.00 | 25.83 | B | C |
| ATOM | 2393 | CB | ALA | B | 84 | 47.207 | -3.776 | 16.470 | 1.00 | 24.41 | B | C |
| ATOM | 2394 | C | ALA | B | 84 | 45.703 | -1.772 | 16.297 | 1.00 | 26.27 | B | C |
| ATOM | 2395 | O | ALA | B | 84 | 44.507 | -1.993 | 16.197 | 1.00 | 24.37 | B | O |
| ATOM | 2396 | N | ASP | B | 85 | 46.184 | -0.769 | 17.027 | 1.00 | 27.62 | B | N |
| ATOM | 2397 | CA | ASP | B | 85 | 45.295 | 0.119 | 17.766 | 1.00 | 30.21 | B | C |
| ATOM | 2398 | CB | ASP | B | 85 | 46.084 | 0.948 | 18.782 | 1.00 | 34.28 | B | C |
| ATOM | 2399 | CG | ASP | B | 85 | 46.619 | 0.109 | 19.927 | 1.00 | 40.65 | B | C |
| ATOM | 2400 | OD1 | ASP | B | 85 | 45.885 | -0.787 | 20.412 | 1.00 | 43.31 | B | O |
| ATOM | 2401 | OD2 | ASP | B | 85 | 47.776 | 0.349 | 20.355 | 1.00 | 45.33 | B | O |
| ATOM | 2402 | C | ASP | B | 85 | 44.436 | 1.058 | 16.907 | 1.00 | 29.53 | B | C |
| ATOM | 2403 | O | ASP | B | 85 | 43.294 | 1.348 | 17.270 | 1.00 | 30.28 | B | O |
| ATOM | 2404 | N | VAL | B | 86 | 44.965 | 1.545 | 15.788 | 1.00 | 28.68 | B | N |
| ATOM | 2405 | CA | VAL | B | 86 | 44.177 | 2.435 | 14.939 | 1.00 | 27.61 | B | C |
| ATOM | 2406 | CB | VAL | B | 86 | 45.064 | 3.321 | 14.048 | 1.00 | 27.16 | B | C |
| ATOM | 2407 | CG1 | VAL | B | 86 | 45.901 | 4.223 | 14.910 | 1.00 | 27.71 | B | C |
| ATOM | 2408 | CG2 | VAL | B | 86 | 45.949 | 2.462 | 13.158 | 1.00 | 26.84 | B | C |
| ATOM | 2409 | C | VAL | B | 86 | 43.214 | 1.658 | 14.054 | 1.00 | 27.48 | B | C |
| ATOM | 2410 | O | VAL | B | 86 | 42.323 | 2.241 | 13.436 | 1.00 | 27.99 | B | O |
| ATOM | 2411 | N | GLY | B | 87 | 43.390 | 0.342 | 13.990 | 1.00 | 27.25 | B | N |
| ATOM | 2412 | CA | GLY | B | 87 | 42.499 | -0.471 | 13.172 | 1.00 | 27.23 | B | C |
| ATOM | 2413 | C | GLY | B | 87 | 43.102 | -1.308 | 12.053 | 1.00 | 27.11 | B | C |
| ATOM | 2414 | O | GLY | B | 87 | 42.366 | -1.923 | 11.291 | 1.00 | 27.58 | B | O |
| ATOM | 2415 | N | CYS | B | 88 | 44.425 | -1.343 | 11.944 | 1.00 | 26.19 | B | N |
| ATOM | 2416 | CA | CYS | B | 88 | 45.071 | -2.130 | 10.903 | 1.00 | 26.54 | B | C |
| ATOM | 2417 | CB | CYS | B | 88 | 46.531 | -1.711 | 10.739 | 1.00 | 26.82 | B | C |
| ATOM | 2418 | SG | CYS | B | 88 | 46.793 | -0.050 | 10.084 | 1.00 | 30.30 | B | S |
| ATOM | 2419 | C | CYS | B | 88 | 45.016 | -3.613 | 11.262 | 1.00 | 25.95 | B | C |
| ATOM | 2420 | O | CYS | B | 88 | 45.102 | -3.981 | 12.430 | 1.00 | 25.20 | B | O |
| ATOM | 2421 | N | SER | B | 89 | 44.887 | -4.461 | 10.251 | 1.00 | 26.81 | B | N |
| ATOM | 2422 | CA | SER | B | 89 | 44.824 | -5.897 | 10.483 | 1.00 | 26.73 | B | C |
| ATOM | 2423 | CB | SER | B | 89 | 43.467 | -6.447 | 10.015 | 1.00 | 27.18 | B | C |
| ATOM | 2424 | OG | SER | B | 89 | 43.207 | -6.122 | 8.656 | 1.00 | 28.77 | B | O |
| ATOM | 2425 | C | SER | B | 89 | 45.961 | -6.655 | 9.797 | 1.00 | 26.65 | B | C |
| ATOM | 2426 | O | SER | B | 89 | 46.172 | -7.833 | 10.082 | 1.00 | 27.15 | B | O |
| ATOM | 2427 | N | MET | B | 90 | 46.690 | -5.977 | 8.909 | 1.00 | 25.25 | B | N |
| ATOM | 2428 | CA | MET | B | 90 | 47.805 | -6.588 | 8.178 | 1.00 | 25.82 | B | C |
| ATOM | 2429 | CB | MET | B | 90 | 47.390 | -6.979 | 6.749 | 1.00 | 24.69 | B | C |
| ATOM | 2430 | CG | MET | B | 90 | 46.014 | -7.603 | 6.551 | 1.00 | 27.70 | B | C |
| ATOM | 2431 | SD | MET | B | 90 | 45.855 | -8.245 | 4.842 | 1.00 | 24.18 | B | S |
| ATOM | 2432 | CE | MET | B | 90 | 46.387 | -9.884 | 5.113 | 1.00 | 26.45 | B | C |
| ATOM | 2433 | C | MET | B | 90 | 48.982 | -5.612 | 8.043 | 1.00 | 25.79 | B | C |
| ATOM | 2434 | O | MET | B | 90 | 48.830 | -4.415 | 8.263 | 1.00 | 27.02 | B | O |
| ATOM | 2435 | N | VAL | B | 91 | 50.148 | -6.134 | 7.666 | 1.00 | 25.33 | B | N |
| ATOM | 2436 | CA | VAL | B | 91 | 51.336 | -5.309 | 7.441 | 1.00 | 23.42 | B | C |
| ATOM | 2437 | CB | VAL | B | 91 | 52.099 | -4.959 | 8.769 | 1.00 | 22.23 | B | C |
| ATOM | 2438 | CG1 | VAL | B | 91 | 52.591 | -6.213 | 9.460 | 1.00 | 21.67 | B | C |
| ATOM | 2439 | CG2 | VAL | B | 91 | 53.255 | -4.034 | 8.465 | 1.00 | 18.12 | B | C |

FIGURE 9 - 37

SUBSTITUTE SHEET (RULE 26)

47/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|----|--------|---------|--------|------|-------|---|---|
| ATOM | 2440 | C | VAL | B | 91 | 52.285 | -6.016 | 6.469 | 1.00 | 23.94 | B | C |
| ATOM | 2441 | O | VAL | B | 91 | 52.504 | -7.223 | 6.559 | 1.00 | 22.79 | B | O |
| ATOM | 2442 | N | LEU | B | 92 | 52.814 | -5.248 | 5.522 | 1.00 | 23.79 | B | N |
| ATOM | 2443 | CA | LEU | B | 92 | 53.737 | -5.766 | 4.527 | 1.00 | 23.98 | B | C |
| ATOM | 2444 | CB | LEU | B | 92 | 53.814 | -4.824 | 3.326 | 1.00 | 22.62 | B | C |
| ATOM | 2445 | CG | LEU | B | 92 | 52.504 | -4.582 | 2.571 | 1.00 | 22.76 | B | C |
| ATOM | 2446 | CD1 | LEU | B | 92 | 52.767 | -3.632 | 1.402 | 1.00 | 20.26 | B | C |
| ATOM | 2447 | CD2 | LEU | B | 92 | 51.932 | -5.913 | 2.075 | 1.00 | 21.47 | B | C |
| ATOM | 2448 | C | LEU | B | 92 | 55.106 | -5.883 | 5.155 | 1.00 | 24.17 | B | C |
| ATOM | 2449 | O | LEU | B | 92 | 55.581 | -4.952 | 5.810 | 1.00 | 25.10 | B | O |
| ATOM | 2450 | N | VAL | B | 93 | 55.748 | -7.023 | 4.959 | 1.00 | 23.36 | B | N |
| ATOM | 2451 | CA | VAL | B | 93 | 57.065 | -7.214 | 5.527 | 1.00 | 23.61 | B | C |
| ATOM | 2452 | CB | VAL | B | 93 | 57.020 | -8.210 | 6.698 | 1.00 | 25.20 | B | C |
| ATOM | 2453 | CG1 | VAL | B | 93 | 58.304 | -8.145 | 7.479 | 1.00 | 28.50 | B | C |
| ATOM | 2454 | CG2 | VAL | B | 93 | 55.862 | -7.876 | 7.607 | 1.00 | 29.55 | B | C |
| ATOM | 2455 | C | VAL | B | 93 | 58.034 | -7.718 | 4.469 | 1.00 | 23.51 | B | C |
| ATOM | 2456 | O | VAL | B | 93 | 57.742 | -8.667 | 3.733 | 1.00 | 21.93 | B | O |
| ATOM | 2457 | N | GLY | B | 94 | 59.180 | -7.057 | 4.381 | 1.00 | 22.72 | B | N |
| ATOM | 2458 | CA | GLY | B | 94 | 60.187 | -7.466 | 3.422 | 1.00 | 24.28 | B | C |
| ATOM | 2459 | C | GLY | B | 94 | 59.965 | -7.105 | 1.965 | 1.00 | 24.47 | B | C |
| ATOM | 2460 | O | GLY | B | 94 | 60.544 | -7.748 | 1.092 | 1.00 | 24.48 | B | O |
| ATOM | 2461 | N | HIS | B | 95 | 59.156 | -6.088 | 1.679 | 1.00 | 24.28 | B | N |
| ATOM | 2462 | CA | HIS | B | 95 | 58.932 | -5.706 | 0.283 | 1.00 | 25.08 | B | C |
| ATOM | 2463 | CB | HIS | B | 95 | 58.143 | -4.400 | 0.202 | 1.00 | 24.85 | B | C |
| ATOM | 2464 | CG | HIS | B | 95 | 57.621 | -4.102 | -1.168 | 1.00 | 26.23 | B | C |
| ATOM | 2465 | CD2 | HIS | B | 95 | 56.443 | -4.416 | -1.759 | 1.00 | 26.78 | B | C |
| ATOM | 2466 | ND1 | HIS | B | 95 | 58.368 | -3.450 | -2.126 | 1.00 | 26.35 | B | N |
| ATOM | 2467 | CE1 | HIS | B | 95 | 57.674 | -3.375 | -3.247 | 1.00 | 26.25 | B | C |
| ATOM | 2468 | NE2 | HIS | B | 95 | 56.502 | -3.954 | -3.052 | 1.00 | 26.92 | B | N |
| ATOM | 2469 | C | HIS | B | 95 | 60.289 | -5.553 | -0.403 | 1.00 | 24.08 | B | C |
| ATOM | 2470 | O | HIS | B | 95 | 61.244 | -5.084 | 0.209 | 1.00 | 23.90 | B | O |
| ATOM | 2471 | N | SER | B | 96 | 60.384 | -5.940 | -1.670 | 1.00 | 24.49 | B | N |
| ATOM | 2472 | CA | SER | B | 96 | 61.668 | -5.862 | -2.370 | 1.00 | 25.33 | B | C |
| ATOM | 2473 | CB | SER | B | 96 | 61.548 | -6.386 | -3.809 | 1.00 | 24.26 | B | C |
| ATOM | 2474 | OG | SER | B | 96 | 60.792 | -5.520 | -4.632 | 1.00 | 26.66 | B | O |
| ATOM | 2475 | C | SER | B | 96 | 62.280 | -4.467 | -2.376 | 1.00 | 25.73 | B | C |
| ATOM | 2476 | O | SER | B | 96 | 63.500 | -4.321 | -2.388 | 1.00 | 24.04 | B | O |
| ATOM | 2477 | N | GLU | B | 97 | 61.438 | -3.438 | -2.374 | 1.00 | 26.85 | B | N |
| ATOM | 2478 | CA | GLU | B | 97 | 61.962 | -2.080 | -2.365 | 1.00 | 26.82 | B | C |
| ATOM | 2479 | CB | GLU | B | 97 | 60.823 | -1.075 | -2.577 | 1.00 | 26.75 | B | C |
| ATOM | 2480 | CG | GLU | B | 97 | 60.407 | -0.941 | -4.039 | 1.00 | 25.34 | B | C |
| ATOM | 2481 | CD | GLU | B | 97 | 58.975 | -0.458 | -4.216 | 1.00 | 27.26 | B | C |
| ATOM | 2482 | OE1 | GLU | B | 97 | 58.431 | 0.170 | -3.277 | 1.00 | 27.65 | B | O |
| ATOM | 2483 | OE2 | GLU | B | 97 | 58.396 | -0.695 | -5.300 | 1.00 | 24.48 | B | O |
| ATOM | 2484 | C | GLU | B | 97 | 62.713 | -1.811 | -1.060 | 1.00 | 25.77 | B | C |
| ATOM | 2485 | O | GLU | B | 97 | 63.750 | -1.165 | -1.055 | 1.00 | 26.01 | B | O |
| ATOM | 2486 | N | ARG | B | 98 | 62.207 | -2.333 | 0.047 | 1.00 | 25.51 | B | N |
| ATOM | 2487 | CA | ARG | B | 98 | 62.871 | -2.122 | 1.322 | 1.00 | 25.91 | B | C |
| ATOM | 2488 | CB | ARG | B | 98 | 61.929 | -2.449 | 2.475 | 1.00 | 23.40 | B | C |
| ATOM | 2489 | CG | ARG | B | 98 | 60.705 | -1.550 | 2.529 | 1.00 | 21.77 | B | C |
| ATOM | 2490 | CD | ARG | B | 98 | 60.006 | -1.658 | 3.867 | 1.00 | 18.48 | B | C |
| ATOM | 2491 | NE | ARG | B | 98 | 60.784 | -1.032 | 4.928 | 1.00 | 18.41 | B | N |
| ATOM | 2492 | CZ | ARG | B | 98 | 60.939 | 0.280 | 5.071 | 1.00 | 21.13 | B | C |
| ATOM | 2493 | NH1 | ARG | B | 98 | 60.363 | 1.123 | 4.222 | 1.00 | 20.27 | B | N |
| ATOM | 2494 | NH2 | ARG | B | 98 | 61.682 | 0.753 | 6.064 | 1.00 | 22.49 | B | N |
| ATOM | 2495 | C | ARG | B | 98 | 64.148 | -2.951 | 1.442 | 1.00 | 28.56 | B | C |
| ATOM | 2496 | O | ARG | B | 98 | 65.095 | -2.554 | 2.134 | 1.00 | 29.22 | B | O |
| ATOM | 2497 | N | ARG | B | 99 | 64.183 | -4.098 | 0.769 | 1.00 | 27.78 | B | N |
| ATOM | 2498 | CA | ARG | B | 99 | 65.361 | -4.943 | 0.833 | 1.00 | 29.08 | B | C |
| ATOM | 2499 | CB | ARG | B | 99 | 65.035 | -6.379 | 0.411 | 1.00 | 28.63 | B | C |
| ATOM | 2500 | CG | ARG | B | 99 | 64.015 | -7.117 | 1.271 | 1.00 | 26.70 | B | C |
| ATOM | 2501 | CD | ARG | B | 99 | 64.040 | -8.594 | 0.896 | 1.00 | 25.52 | B | C |
| ATOM | 2502 | NE | ARG | B | 99 | 62.891 | -9.341 | 1.395 | 1.00 | 24.37 | B | N |
| ATOM | 2503 | CZ | ARG | B | 99 | 62.976 | -10.372 | 2.227 | 1.00 | 22.76 | B | C |
| ATOM | 2504 | NH1 | ARG | B | 99 | 64.157 | -10.780 | 2.670 | 1.00 | 21.78 | B | N |
| ATOM | 2505 | NH2 | ARG | B | 99 | 61.881 | -11.017 | 2.589 | 1.00 | 23.44 | B | N |

FIGURE 9- 38

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|--------|------|-------|---|---|
| ATOM | 2506 | C | ARG | B | 99 | 66.484 | -4.425 | -0.061 | 1.00 | 30.40 | B | C |
| ATOM | 2507 | O | ARG | B | 99 | 67.628 | -4.328 | 0.377 | 1.00 | 30.42 | B | O |
| ATOM | 2508 | N | LEU | B | 100 | 66.144 | -4.079 | -1.303 | 1.00 | 31.24 | B | N |
| ATOM | 2509 | CA | LEU | B | 100 | 67.130 | -3.635 | -2.290 | 1.00 | 32.17 | B | C |
| ATOM | 2510 | CB | LEU | B | 100 | 66.655 | -4.013 | -3.698 | 1.00 | 33.44 | B | C |
| ATOM | 2511 | CG | LEU | B | 100 | 66.217 | -5.459 | -3.949 | 1.00 | 34.34 | B | C |
| ATOM | 2512 | CD1 | LEU | B | 100 | 65.988 | -5.628 | -5.428 | 1.00 | 35.69 | B | C |
| ATOM | 2513 | CD2 | LEU | B | 100 | 67.274 | -6.442 | -3.476 | 1.00 | 35.35 | B | C |
| ATOM | 2514 | C | LEU | B | 100 | 67.498 | -2.162 | -2.296 | 1.00 | 33.08 | B | C |
| ATOM | 2515 | O | LEU | B | 100 | 68.642 | -1.801 | -2.584 | 1.00 | 33.91 | B | O |
| ATOM | 2516 | N | ILE | B | 101 | 66.532 | -1.305 | -2.006 | 1.00 | 33.00 | B | N |
| ATOM | 2517 | CA | ILE | B | 101 | 66.792 | 0.119 | -2.019 | 1.00 | 32.71 | B | C |
| ATOM | 2518 | CB | ILE | B | 101 | 65.565 | 0.897 | -2.561 | 1.00 | 31.87 | B | C |
| ATOM | 2519 | CG2 | ILE | B | 101 | 65.877 | 2.378 | -2.636 | 1.00 | 32.37 | B | C |
| ATOM | 2520 | CG1 | ILE | B | 101 | 65.219 | 0.387 | -3.966 | 1.00 | 34.11 | B | C |
| ATOM | 2521 | CD1 | ILE | B | 101 | 63.967 | 0.994 | -4.573 | 1.00 | 34.50 | B | C |
| ATOM | 2522 | C | ILE | B | 101 | 67.161 | 0.641 | -0.643 | 1.00 | 32.97 | B | C |
| ATOM | 2523 | O | ILE | B | 101 | 67.992 | 1.541 | -0.525 | 1.00 | 33.57 | B | O |
| ATOM | 2524 | N | LEU | B | 102 | 66.559 | 0.061 | 0.393 | 1.00 | 32.60 | B | N |
| ATOM | 2525 | CA | LEU | B | 102 | 66.808 | 0.498 | 1.758 | 1.00 | 32.93 | B | C |
| ATOM | 2526 | CB | LEU | B | 102 | 65.480 | 0.645 | 2.506 | 1.00 | 32.48 | B | C |
| ATOM | 2527 | CG | LEU | B | 102 | 64.450 | 1.554 | 1.820 | 1.00 | 33.34 | B | C |
| ATOM | 2528 | CD1 | LEU | B | 102 | 63.193 | 1.629 | 2.667 | 1.00 | 32.88 | B | C |
| ATOM | 2529 | CD2 | LEU | B | 102 | 65.036 | 2.949 | 1.609 | 1.00 | 33.61 | B | C |
| ATOM | 2530 | C | LEU | B | 102 | 67.756 | -0.385 | 2.557 | 1.00 | 32.70 | B | C |
| ATOM | 2531 | O | LEU | B | 102 | 68.178 | -0.001 | 3.644 | 1.00 | 32.41 | B | O |
| ATOM | 2532 | N | GLY | B | 103 | 68.074 | -1.568 | 2.030 | 1.00 | 32.88 | B | N |
| ATOM | 2533 | CA | GLY | B | 103 | 69.002 | -2.464 | 2.707 | 1.00 | 32.27 | B | C |
| ATOM | 2534 | C | GLY | B | 103 | 68.496 | -3.377 | 3.817 | 1.00 | 33.21 | B | C |
| ATOM | 2535 | O | GLY | B | 103 | 69.298 | -3.917 | 4.579 | 1.00 | 32.63 | B | O |
| ATOM | 2536 | N | GLU | B | 104 | 67.388 | -3.583 | 3.922 | 1.00 | 32.89 | B | N |
| ATOM | 2537 | CA | GLU | B | 104 | 66.678 | -4.444 | 4.983 | 1.00 | 32.85 | B | C |
| ATOM | 2538 | CB | GLU | B | 104 | 65.158 | -4.341 | 5.074 | 1.00 | 34.28 | B | C |
| ATOM | 2539 | CG | GLU | B | 104 | 64.671 | -2.921 | 5.225 | 1.00 | 35.28 | B | C |
| ATOM | 2540 | CD | GLU | B | 104 | 63.479 | -2.801 | 6.144 | 1.00 | 34.52 | B | C |
| ATOM | 2541 | OE1 | GLU | B | 104 | 62.545 | -3.622 | 6.032 | 1.00 | 31.46 | B | O |
| ATOM | 2542 | OE2 | GLU | B | 104 | 63.485 | -1.864 | 6.976 | 1.00 | 37.20 | B | O |
| ATOM | 2543 | C | GLU | B | 104 | 67.097 | -5.898 | 4.780 | 1.00 | 32.53 | B | C |
| ATOM | 2544 | O | GLU | B | 104 | 66.818 | -6.501 | 3.740 | 1.00 | 32.73 | B | O |
| ATOM | 2545 | N | SER | B | 105 | 67.766 | -6.455 | 5.786 | 1.00 | 31.46 | B | N |
| ATOM | 2546 | CA | SER | B | 105 | 68.251 | -7.825 | 5.717 | 1.00 | 31.24 | B | C |
| ATOM | 2547 | CB | SER | B | 105 | 69.534 | -7.953 | 6.543 | 1.00 | 31.81 | B | C |
| ATOM | 2548 | OG | SER | B | 105 | 69.296 | -7.616 | 7.897 | 1.00 | 32.78 | B | O |
| ATOM | 2549 | C | SER | B | 105 | 67.211 | -8.826 | 6.210 | 1.00 | 31.41 | B | C |
| ATOM | 2550 | O | SER | B | 105 | 66.206 | -8.446 | 6.822 | 1.00 | 32.05 | B | O |
| ATOM | 2551 | N | ASP | B | 106 | 67.454 | -10.107 | 5.941 | 1.00 | 30.19 | B | N |
| ATOM | 2552 | CA | ASP | B | 106 | 66.534 | -11.159 | 6.365 | 1.00 | 31.36 | B | C |
| ATOM | 2553 | CB | ASP | B | 106 | 67.069 | -12.535 | 5.961 | 1.00 | 33.09 | B | C |
| ATOM | 2554 | CG | ASP | B | 106 | 67.431 | -12.613 | 4.483 | 1.00 | 35.40 | B | C |
| ATOM | 2555 | OD1 | ASP | B | 106 | 66.791 | -11.908 | 3.678 | 1.00 | 35.72 | B | O |
| ATOM | 2556 | OD2 | ASP | B | 106 | 68.341 | -13.396 | 4.124 | 1.00 | 37.24 | B | O |
| ATOM | 2557 | C | ASP | B | 106 | 66.355 | -11.111 | 7.878 | 1.00 | 30.43 | B | C |
| ATOM | 2558 | O | ASP | B | 106 | 65.303 | -11.464 | 8.408 | 1.00 | 29.51 | B | O |
| ATOM | 2559 | N | GLU | B | 107 | 67.405 | -10.658 | 8.553 | 1.00 | 31.76 | B | N |
| ATOM | 2560 | CA | GLU | B | 107 | 67.448 | -10.530 | 10.009 | 1.00 | 31.30 | B | C |
| ATOM | 2561 | CB | GLU | B | 107 | 68.826 | -9.997 | 10.409 | 1.00 | 34.55 | B | C |
| ATOM | 2562 | CG | GLU | B | 107 | 69.405 | -10.534 | 11.696 | 1.00 | 41.41 | B | C |
| ATOM | 2563 | CD | GLU | B | 107 | 70.888 | -10.177 | 11.845 | 1.00 | 46.02 | B | C |
| ATOM | 2564 | OE1 | GLU | B | 107 | 71.563 | -10.756 | 12.735 | 1.00 | 47.44 | B | O |
| ATOM | 2565 | OE2 | GLU | B | 107 | 71.377 | -9.316 | 11.068 | 1.00 | 48.01 | B | O |
| ATOM | 2566 | C | GLU | B | 107 | 66.364 | -9.548 | 10.440 | 1.00 | 29.88 | B | C |
| ATOM | 2567 | O | GLU | B | 107 | 65.525 | -9.853 | 11.284 | 1.00 | 29.77 | B | O |
| ATOM | 2568 | N | VAL | B | 108 | 66.403 | -8.362 | 9.847 | 1.00 | 29.22 | B | N |
| ATOM | 2569 | CA | VAL | B | 108 | 65.444 | -7.305 | 10.140 | 1.00 | 27.29 | B | C |
| ATOM | 2570 | CB | VAL | B | 108 | 65.838 | -6.009 | 9.409 | 1.00 | 27.08 | B | C |
| ATOM | 2571 | CG1 | VAL | B | 108 | 64.731 | -4.969 | 9.533 | 1.00 | 24.41 | B | C |

FIGURE 9- 39

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | |
|------|------|---------------|--------|---------|--------|------|-------|---|---|
| ATOM | 2572 | CG2 VAL B 108 | 67.152 | -5.480 | 9.992 | 1.00 | 27.41 | B | C |
| ATOM | 2573 | C VAL B 108 | 64.032 | -7.707 | 9.733 | 1.00 | 26.82 | B | C |
| ATOM | 2574 | O VAL B 108 | 63.088 | -7.524 | 10.497 | 1.00 | 26.60 | B | O |
| ATOM | 2575 | N VAL B 109 | 63.886 | -8.253 | 8.530 | 1.00 | 26.25 | B | N |
| ATOM | 2576 | CA VAL B 109 | 62.574 | -8.673 | 8.058 | 1.00 | 24.93 | B | C |
| ATOM | 2577 | CB VAL B 109 | 62.658 | -9.355 | 6.666 | 1.00 | 24.14 | B | C |
| ATOM | 2578 | CG1 VAL B 109 | 61.339 | -10.051 | 6.337 | 1.00 | 26.04 | B | C |
| ATOM | 2579 | CG2 VAL B 109 | 62.960 | -8.318 | 5.593 | 1.00 | 24.79 | B | C |
| ATOM | 2580 | C VAL B 109 | 61.964 | -9.655 | 9.047 | 1.00 | 25.08 | B | C |
| ATOM | 2581 | O VAL B 109 | 60.793 | -9.559 | 9.394 | 1.00 | 24.13 | B | O |
| ATOM | 2582 | N SER B 110 | 62.769 | -10.599 | 9.508 | 1.00 | 24.84 | B | N |
| ATOM | 2583 | CA SER B 110 | 62.266 | -11.589 | 10.429 | 1.00 | 27.07 | B | C |
| ATOM | 2584 | CB SER B 110 | 63.326 | -12.640 | 10.713 | 1.00 | 28.86 | B | C |
| ATOM | 2585 | OG SER B 110 | 62.726 | -13.729 | 11.378 | 1.00 | 32.52 | B | O |
| ATOM | 2586 | C SER B 110 | 61.803 | -10.967 | 11.733 | 1.00 | 27.23 | B | C |
| ATOM | 2587 | O SER B 110 | 60.721 | -11.288 | 12.223 | 1.00 | 26.48 | B | O |
| ATOM | 2588 | N ARG B 111 | 62.624 | -10.087 | 12.302 | 1.00 | 26.92 | B | N |
| ATOM | 2589 | CA ARG B 111 | 62.243 | -9.434 | 13.543 | 1.00 | 27.50 | B | C |
| ATOM | 2590 | CB ARG B 111 | 63.314 | -8.435 | 13.995 | 1.00 | 30.02 | B | C |
| ATOM | 2591 | CG ARG B 111 | 64.340 | -9.041 | 14.937 | 1.00 | 33.73 | B | C |
| ATOM | 2592 | CD ARG B 111 | 64.751 | -8.042 | 16.024 | 1.00 | 37.05 | B | C |
| ATOM | 2593 | NE ARG B 111 | 65.698 | -7.046 | 15.532 | 1.00 | 39.08 | B | N |
| ATOM | 2594 | CZ ARG B 111 | 65.981 | -5.907 | 16.161 | 1.00 | 40.34 | B | C |
| ATOM | 2595 | NH1 ARG B 111 | 65.384 | -5.616 | 17.315 | 1.00 | 40.62 | B | N |
| ATOM | 2596 | NH2 ARG B 111 | 66.851 | -5.050 | 15.625 | 1.00 | 39.48 | B | N |
| ATOM | 2597 | C ARG B 111 | 60.911 | -8.715 | 13.376 | 1.00 | 25.86 | B | C |
| ATOM | 2598 | O ARG B 111 | 60.084 | -8.734 | 14.280 | 1.00 | 25.33 | B | O |
| ATOM | 2599 | N LYS B 112 | 60.701 | -8.100 | 12.214 | 1.00 | 24.77 | B | N |
| ATOM | 2600 | CA LYS B 112 | 59.462 | -7.372 | 11.948 | 1.00 | 25.39 | B | C |
| ATOM | 2601 | CB LYS B 112 | 59.587 | -6.547 | 10.654 | 1.00 | 25.43 | B | C |
| ATOM | 2602 | CG LYS B 112 | 60.587 | -5.391 | 10.752 | 1.00 | 26.37 | B | C |
| ATOM | 2603 | CD LYS B 112 | 60.548 | -4.479 | 9.522 | 1.00 | 25.10 | B | C |
| ATOM | 2604 | CE LYS B 112 | 61.430 | -3.250 | 9.731 | 1.00 | 26.49 | B | C |
| ATOM | 2605 | NZ LYS B 112 | 61.392 | -2.328 | 8.564 | 1.00 | 25.15 | B | N |
| ATOM | 2606 | C LYS B 112 | 58.272 | -8.315 | 11.849 | 1.00 | 24.14 | B | C |
| ATOM | 2607 | O LYS B 112 | 57.190 | -8.031 | 12.363 | 1.00 | 23.48 | B | O |
| ATOM | 2608 | N PHE B 113 | 58.473 | -9.435 | 11.173 | 1.00 | 23.74 | B | N |
| ATOM | 2609 | CA PHE B 113 | 57.418 | -10.426 | 11.027 | 1.00 | 24.41 | B | C |
| ATOM | 2610 | CB PHE B 113 | 57.974 | -11.643 | 10.274 | 1.00 | 22.63 | B | C |
| ATOM | 2611 | CG PHE B 113 | 56.960 | -12.726 | 9.989 | 1.00 | 22.59 | B | C |
| ATOM | 2612 | CD1 PHE B 113 | 56.378 | -13.449 | 11.019 | 1.00 | 20.86 | B | C |
| ATOM | 2613 | CD2 PHE B 113 | 56.665 | -13.084 | 8.671 | 1.00 | 23.17 | B | C |
| ATOM | 2614 | CE1 PHE B 113 | 55.526 | -14.521 | 10.744 | 1.00 | 22.73 | B | C |
| ATOM | 2615 | CE2 PHE B 113 | 55.814 | -14.152 | 8.389 | 1.00 | 21.80 | B | C |
| ATOM | 2616 | CZ PHE B 113 | 55.247 | -14.872 | 9.427 | 1.00 | 19.67 | B | C |
| ATOM | 2617 | C PHE B 113 | 56.973 | -10.803 | 12.443 | 1.00 | 23.46 | B | C |
| ATOM | 2618 | O PHE B 113 | 55.798 | -10.710 | 12.783 | 1.00 | 23.49 | B | O |
| ATOM | 2619 | N ALA B 114 | 57.931 | -11.203 | 13.272 | 1.00 | 24.60 | B | N |
| ATOM | 2620 | CA ALA B 114 | 57.651 | -11.592 | 14.651 | 1.00 | 24.80 | B | C |
| ATOM | 2621 | CB ALA B 114 | 58.936 | -12.005 | 15.335 | 1.00 | 24.80 | B | C |
| ATOM | 2622 | C ALA B 114 | 56.969 | -10.473 | 15.448 | 1.00 | 24.78 | B | C |
| ATOM | 2623 | O ALA B 114 | 55.952 | -10.695 | 16.101 | 1.00 | 25.03 | B | O |
| ATOM | 2624 | N ALA B 115 | 57.530 | -9.273 | 15.388 | 1.00 | 24.89 | B | N |
| ATOM | 2625 | CA ALA B 115 | 56.962 | -8.151 | 16.114 | 1.00 | 26.23 | B | C |
| ATOM | 2626 | CB ALA B 115 | 57.800 | -6.893 | 15.873 | 1.00 | 24.51 | B | C |
| ATOM | 2627 | C ALA B 115 | 55.505 | -7.921 | 15.692 | 1.00 | 26.26 | B | C |
| ATOM | 2628 | O ALA B 115 | 54.618 | -7.764 | 16.532 | 1.00 | 28.67 | B | O |
| ATOM | 2629 | N ALA B 116 | 55.259 | -7.915 | 14.389 | 1.00 | 26.25 | B | N |
| ATOM | 2630 | CA ALA B 116 | 53.913 | -7.707 | 13.876 | 1.00 | 25.42 | B | C |
| ATOM | 2631 | CB ALA B 116 | 53.937 | -7.720 | 12.345 | 1.00 | 25.11 | B | C |
| ATOM | 2632 | C ALA B 116 | 52.916 | -8.749 | 14.403 | 1.00 | 26.50 | B | C |
| ATOM | 2633 | O ALA B 116 | 51.810 | -8.398 | 14.815 | 1.00 | 26.19 | B | O |
| ATOM | 2634 | N GLN B 117 | 53.310 | -10.022 | 14.399 | 1.00 | 26.20 | B | N |
| ATOM | 2635 | CA GLN B 117 | 52.431 | -11.101 | 14.861 | 1.00 | 27.92 | B | C |
| ATOM | 2636 | CB GLN B 117 | 53.020 | -12.474 | 14.514 | 1.00 | 28.38 | B | C |
| ATOM | 2637 | CG GLN B 117 | 53.068 | -12.785 | 13.035 | 1.00 | 29.14 | B | C |

FIGURE 9 - 40

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|--------|------|-------|---|---|
| ATOM | 2638 | CD | GLN | B | 117 | 53.069 | -14.271 | 12.777 | 1.00 | 30.51 | B | C |
| ATOM | 2639 | OE1 | GLN | B | 117 | 53.894 | -15.011 | 13.320 | 1.00 | 30.34 | B | O |
| ATOM | 2640 | NE2 | GLN | B | 117 | 52.138 | -14.724 | 11.948 | 1.00 | 31.29 | B | N |
| ATOM | 2641 | C | GLN | B | 117 | 52.117 | -11.076 | 16.356 | 1.00 | 28.27 | B | C |
| ATOM | 2642 | O | GLN | B | 117 | 51.009 | -11.424 | 16.766 | 1.00 | 26.72 | B | O |
| ATOM | 2643 | N | SER | B | 118 | 53.089 | -10.696 | 17.178 | 1.00 | 27.52 | B | N |
| ATOM | 2644 | CA | SER | B | 118 | 52.832 | -10.650 | 18.609 | 1.00 | 29.13 | B | C |
| ATOM | 2645 | CB | SER | B | 118 | 54.130 | -10.409 | 19.384 | 1.00 | 30.66 | B | C |
| ATOM | 2646 | OG | SER | B | 118 | 54.631 | -9.115 | 19.114 | 1.00 | 33.25 | B | O |
| ATOM | 2647 | C | SER | B | 118 | 51.819 | -9.543 | 18.923 | 1.00 | 28.89 | B | C |
| ATOM | 2648 | O | SER | B | 118 | 51.335 | -9.447 | 20.041 | 1.00 | 29.05 | B | O |
| ATOM | 2649 | N | CYS | B | 119 | 51.507 | -8.704 | 17.936 | 1.00 | 29.05 | B | N |
| ATOM | 2650 | CA | CYS | B | 119 | 50.529 | -7.633 | 18.133 | 1.00 | 28.47 | B | C |
| ATOM | 2651 | CB | CYS | B | 119 | 51.063 | -6.298 | 17.589 | 1.00 | 29.79 | B | C |
| ATOM | 2652 | SG | CYS | B | 119 | 52.378 | -5.530 | 18.608 | 1.00 | 31.56 | B | S |
| ATOM | 2653 | C | CYS | B | 119 | 49.201 | -7.983 | 17.458 | 1.00 | 28.08 | B | C |
| ATOM | 2654 | O | CYS | B | 119 | 48.285 | -7.163 | 17.401 | 1.00 | 28.18 | B | O |
| ATOM | 2655 | N | GLY | B | 120 | 49.103 | -9.203 | 16.943 | 1.00 | 27.02 | B | N |
| ATOM | 2656 | CA | GLY | B | 120 | 47.874 | -9.633 | 16.300 | 1.00 | 27.65 | B | C |
| ATOM | 2657 | C | GLY | B | 120 | 47.719 | -9.271 | 14.832 | 1.00 | 28.01 | B | C |
| ATOM | 2658 | O | GLY | B | 120 | 46.670 | -9.533 | 14.241 | 1.00 | 28.97 | B | O |
| ATOM | 2659 | N | LEU | B | 121 | 48.746 | -8.674 | 14.236 | 1.00 | 26.53 | B | N |
| ATOM | 2660 | CA | LEU | B | 121 | 48.672 | -8.298 | 12.832 | 1.00 | 25.13 | B | C |
| ATOM | 2661 | CB | LEU | B | 121 | 49.608 | -7.125 | 12.535 | 1.00 | 23.72 | B | C |
| ATOM | 2662 | CG | LEU | B | 121 | 49.363 | -5.788 | 13.235 | 1.00 | 26.83 | B | C |
| ATOM | 2663 | CD1 | LEU | B | 121 | 50.407 | -4.766 | 12.752 | 1.00 | 24.68 | B | C |
| ATOM | 2664 | CD2 | LEU | B | 121 | 47.958 | -5.292 | 12.929 | 1.00 | 24.19 | B | C |
| ATOM | 2665 | C | LEU | B | 121 | 49.070 | -9.478 | 11.963 | 1.00 | 24.80 | B | C |
| ATOM | 2666 | O | LEU | B | 121 | 49.839 | -10.343 | 12.387 | 1.00 | 23.22 | B | O |
| ATOM | 2667 | N | VAL | B | 122 | 48.538 | -9.514 | 10.748 | 1.00 | 24.46 | B | N |
| ATOM | 2668 | CA | VAL | B | 122 | 48.855 | -10.580 | 9.812 | 1.00 | 23.86 | B | C |
| ATOM | 2669 | CB | VAL | B | 122 | 47.596 | -11.012 | 9.019 | 1.00 | 23.77 | B | C |
| ATOM | 2670 | CG1 | VAL | B | 122 | 47.978 | -11.930 | 7.862 | 1.00 | 23.35 | B | C |
| ATOM | 2671 | CG2 | VAL | B | 122 | 46.643 | -11.742 | 9.954 | 1.00 | 26.09 | B | C |
| ATOM | 2672 | O | VAL | B | 122 | 49.933 | -10.052 | 8.866 | 1.00 | 24.11 | B | C |
| ATOM | 2673 | O | VAL | B | 122 | 49.677 | -9.184 | 8.025 | 1.00 | 24.01 | B | O |
| ATOM | 2674 | N | PRO | B | 123 | 51.162 | -10.561 | 8.996 | 1.00 | 22.93 | B | N |
| ATOM | 2675 | CD | PRO | B | 123 | 51.740 | -11.471 | 10.002 | 1.00 | 23.87 | B | C |
| ATOM | 2676 | CA | PRO | B | 123 | 52.195 | -10.056 | 8.100 | 1.00 | 23.39 | B | C |
| ATOM | 2677 | CB | PRO | B | 123 | 53.475 | -10.384 | 8.853 | 1.00 | 23.77 | B | C |
| ATOM | 2678 | CG | PRO | B | 123 | 53.144 | -11.706 | 9.466 | 1.00 | 22.10 | B | C |
| ATOM | 2679 | C | PRO | B | 123 | 52.157 | -10.695 | 6.724 | 1.00 | 24.24 | B | C |
| ATOM | 2680 | O | PRO | B | 123 | 51.905 | -11.894 | 6.591 | 1.00 | 25.37 | B | O |
| ATOM | 2681 | N | VAL | B | 124 | 52.387 | -9.869 | 5.709 | 1.00 | 24.10 | B | N |
| ATOM | 2682 | CA | VAL | B | 124 | 52.446 | -10.311 | 4.323 | 1.00 | 24.06 | B | C |
| ATOM | 2683 | CB | VAL | B | 124 | 51.757 | -9.306 | 3.359 | 1.00 | 22.66 | B | C |
| ATOM | 2684 | CG1 | VAL | B | 124 | 52.010 | -9.720 | 1.920 | 1.00 | 23.24 | B | C |
| ATOM | 2685 | CG2 | VAL | B | 124 | 50.266 | -9.246 | 3.631 | 1.00 | 22.62 | B | C |
| ATOM | 2686 | C | VAL | B | 124 | 53.937 | -10.361 | 3.982 | 1.00 | 24.56 | B | C |
| ATOM | 2687 | O | VAL | B | 124 | 54.533 | -9.342 | 3.632 | 1.00 | 23.15 | B | O |
| ATOM | 2688 | N | LEU | B | 125 | 54.534 | -11.541 | 4.110 | 1.00 | 25.07 | B | N |
| ATOM | 2689 | CA | LEU | B | 125 | 55.953 | -11.713 | 3.825 | 1.00 | 24.63 | B | C |
| ATOM | 2690 | CB | LEU | B | 125 | 56.466 | -13.006 | 4.453 | 1.00 | 24.91 | B | C |
| ATOM | 2691 | CG | LEU | B | 125 | 57.920 | -13.349 | 4.125 | 1.00 | 24.33 | B | C |
| ATOM | 2692 | CD1 | LEU | B | 125 | 58.848 | -12.309 | 4.751 | 1.00 | 25.50 | B | C |
| ATOM | 2693 | CD2 | LEU | B | 125 | 58.237 | -14.747 | 4.659 | 1.00 | 25.12 | B | C |
| ATOM | 2694 | C | LEU | B | 125 | 56.234 | -11.738 | 2.338 | 1.00 | 23.94 | B | C |
| ATOM | 2695 | O | LEU | B | 125 | 55.786 | -12.644 | 1.629 | 1.00 | 24.25 | B | O |
| ATOM | 2696 | N | CYS | B | 126 | 56.979 | -10.740 | 1.874 | 1.00 | 23.36 | B | N |
| ATOM | 2697 | CA | CYS | B | 126 | 57.340 | -10.615 | 0.464 | 1.00 | 23.33 | B | C |
| ATOM | 2698 | CB | CYS | B | 126 | 57.399 | -9.135 | 0.057 | 1.00 | 24.70 | B | C |
| ATOM | 2699 | SG | CYS | B | 126 | 55.837 | -8.217 | 0.213 | 1.00 | 28.49 | B | S |
| ATOM | 2700 | C | CYS | B | 126 | 58.691 | -11.269 | 0.168 | 1.00 | 23.06 | B | C |
| ATOM | 2701 | O | CYS | B | 126 | 59.669 | -11.031 | 0.867 | 1.00 | 20.87 | B | O |
| ATOM | 2702 | N | VAL | B | 127 | 58.735 | -12.079 | -0.886 | 1.00 | 23.14 | B | N |
| ATOM | 2703 | CA | VAL | B | 127 | 59.954 | -12.771 | -1.298 | 1.00 | 22.72 | B | C |

FIGURE 9 - 41

SUBSTITUTE SHEET (RULE 26)

51/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|---------|------|-------|---|---|
| ATOM | 2704 | CB | VAL | B | 127 | 59.889 | -14.278 | -0.954 | 1.00 | 22.50 | B | C |
| ATOM | 2705 | CG1 | VAL | B | 127 | 59.762 | -14.451 | 0.540 | 1.00 | 22.33 | B | C |
| ATOM | 2706 | CG2 | VAL | B | 127 | 58.705 | -14.932 | -1.652 | 1.00 | 19.85 | B | C |
| ATOM | 2707 | C | VAL | B | 127 | 60.104 | -12.611 | -2.808 | 1.00 | 24.27 | B | C |
| ATOM | 2708 | O | VAL | B | 127 | 59.113 | -12.422 | -3.507 | 1.00 | 24.99 | B | O |
| ATOM | 2709 | N | GLY | B | 128 | 61.335 | -12.676 | -3.308 | 1.00 | 25.84 | B | N |
| ATOM | 2710 | CA | GLY | B | 128 | 61.540 | -12.533 | -4.740 | 1.00 | 28.99 | B | C |
| ATOM | 2711 | C | GLY | B | 128 | 62.991 | -12.400 | -5.156 | 1.00 | 30.79 | B | C |
| ATOM | 2712 | O | GLY | B | 128 | 63.832 | -11.970 | -4.365 | 1.00 | 30.59 | B | O |
| ATOM | 2713 | N | GLU | B | 129 | 63.281 | -12.774 | -6.401 | 1.00 | 33.20 | B | N |
| ATOM | 2714 | CA | GLU | B | 129 | 64.637 | -12.706 | -6.938 | 1.00 | 34.31 | B | C |
| ATOM | 2715 | CB | GLU | B | 129 | 65.041 | -14.036 | -7.576 | 1.00 | 36.36 | B | C |
| ATOM | 2716 | CG | GLU | B | 129 | 64.570 | -14.182 | -9.017 | 1.00 | 38.22 | B | C |
| ATOM | 2717 | CD | GLU | B | 129 | 63.131 | -14.661 | -9.139 | 1.00 | 40.90 | B | C |
| ATOM | 2718 | OE1 | GLU | B | 129 | 62.270 | -14.281 | -8.308 | 1.00 | 40.51 | B | O |
| ATOM | 2719 | OE2 | GLU | B | 129 | 62.861 | -15.415 | -10.092 | 1.00 | 42.53 | B | O |
| ATOM | 2720 | C | GLU | B | 129 | 64.748 | -11.621 | -8.001 | 1.00 | 35.90 | B | C |
| ATOM | 2721 | O | GLU | B | 129 | 63.798 | -11.372 | -8.755 | 1.00 | 35.20 | B | O |
| ATOM | 2722 | N | THR | B | 130 | 65.921 | -10.998 | -8.067 | 1.00 | 36.18 | B | N |
| ATOM | 2723 | CA | THR | B | 130 | 66.197 | -9.944 | -9.035 | 1.00 | 38.25 | B | C |
| ATOM | 2724 | CB | THR | B | 130 | 67.409 | -9.105 | -8.602 | 1.00 | 38.49 | B | C |
| ATOM | 2725 | OG1 | THR | B | 130 | 68.596 | -9.897 | -8.724 | 1.00 | 37.57 | B | O |
| ATOM | 2726 | CG2 | THR | B | 130 | 67.269 | -8.682 | -7.147 | 1.00 | 37.17 | B | C |
| ATOM | 2727 | C | THR | B | 130 | 66.514 | -10.564 | -10.394 | 1.00 | 40.03 | B | C |
| ATOM | 2728 | O | THR | B | 130 | 66.758 | -11.772 | -10.498 | 1.00 | 40.35 | B | O |
| ATOM | 2729 | N | ARG | B | 131 | 66.510 | -9.732 | -11.431 | 1.00 | 41.20 | B | N |
| ATOM | 2730 | CA | ARG | B | 131 | 66.790 | -10.179 | -12.792 | 1.00 | 43.29 | B | C |
| ATOM | 2731 | CB | ARG | B | 131 | 66.849 | -8.972 | -13.740 | 1.00 | 46.06 | B | C |
| ATOM | 2732 | CG | ARG | B | 131 | 67.201 | -9.308 | -15.193 | 1.00 | 49.17 | B | C |
| ATOM | 2733 | CD | ARG | B | 131 | 67.917 | -8.116 | -15.860 | 1.00 | 52.24 | B | C |
| ATOM | 2734 | NE | ARG | B | 131 | 68.325 | -8.411 | -17.231 | 1.00 | 54.48 | B | N |
| ATOM | 2735 | CZ | ARG | B | 131 | 69.169 | -7.662 | -17.944 | 1.00 | 56.08 | B | C |
| ATOM | 2736 | NH1 | ARG | B | 131 | 69.707 | -6.564 | -17.420 | 1.00 | 55.93 | B | N |
| ATOM | 2737 | NH2 | ARG | B | 131 | 69.474 | -8.010 | -19.191 | 1.00 | 56.81 | B | N |
| ATOM | 2738 | C | ARG | B | 131 | 68.102 | -10.942 | -12.875 | 1.00 | 42.45 | B | C |
| ATOM | 2739 | O | ARG | B | 131 | 68.240 | -11.857 | -13.681 | 1.00 | 42.37 | B | O |
| ATOM | 2740 | N | ALA | B | 132 | 69.055 | -10.558 | -12.033 | 1.00 | 42.53 | B | N |
| ATOM | 2741 | CA | ALA | B | 132 | 70.374 | -11.173 | -12.008 | 1.00 | 43.22 | B | C |
| ATOM | 2742 | CB | ALA | B | 132 | 71.332 | -10.295 | -11.215 | 1.00 | 43.49 | B | C |
| ATOM | 2743 | C | ALA | B | 132 | 70.366 | -12.570 | -11.415 | 1.00 | 44.09 | B | C |
| ATOM | 2744 | O | ALA | B | 132 | 71.071 | -13.465 | -11.887 | 1.00 | 43.58 | B | O |
| ATOM | 2745 | N | GLU | B | 133 | 69.582 | -12.746 | -10.359 | 1.00 | 44.51 | B | N |
| ATOM | 2746 | CA | GLU | B | 133 | 69.493 | -14.037 | -9.693 | 1.00 | 44.86 | B | C |
| ATOM | 2747 | CB | GLU | B | 133 | 68.850 | -13.858 | -8.304 | 1.00 | 44.26 | B | C |
| ATOM | 2748 | CG | GLU | B | 133 | 69.542 | -12.790 | -7.445 | 1.00 | 44.26 | B | C |
| ATOM | 2749 | CD | GLU | B | 133 | 68.815 | -12.489 | -6.133 | 1.00 | 44.79 | B | C |
| ATOM | 2750 | OE1 | GLU | B | 133 | 67.604 | -12.166 | -6.182 | 1.00 | 43.17 | B | O |
| ATOM | 2751 | OE2 | GLU | B | 133 | 69.463 | -12.561 | -5.054 | 1.00 | 43.65 | B | O |
| ATOM | 2752 | C | GLU | B | 133 | 68.705 | -15.045 | -10.533 | 1.00 | 44.82 | B | C |
| ATOM | 2753 | O | GLU | B | 133 | 68.941 | -16.251 | -10.449 | 1.00 | 44.34 | B | O |
| ATOM | 2754 | N | ARG | B | 134 | 67.773 | -14.558 | -11.346 | 1.00 | 46.19 | B | N |
| ATOM | 2755 | CA | ARG | B | 134 | 66.980 | -15.458 | -12.184 | 1.00 | 48.45 | B | C |
| ATOM | 2756 | CB | ARG | B | 134 | 65.745 | -14.751 | -12.732 | 1.00 | 48.08 | B | C |
| ATOM | 2757 | CG | ARG | B | 134 | 64.800 | -15.683 | -13.461 | 1.00 | 49.27 | B | C |
| ATOM | 2758 | CD | ARG | B | 134 | 63.567 | -14.950 | -13.954 | 1.00 | 51.86 | B | C |
| ATOM | 2759 | NE | ARG | B | 134 | 62.506 | -15.885 | -14.334 | 1.00 | 54.78 | B | N |
| ATOM | 2760 | CZ | ARG | B | 134 | 61.843 | -16.653 | -13.469 | 1.00 | 56.47 | B | C |
| ATOM | 2761 | NH1 | ARG | B | 134 | 62.130 | -16.596 | -12.170 | 1.00 | 56.59 | B | N |
| ATOM | 2762 | NH2 | ARG | B | 134 | 60.898 | -17.490 | -13.903 | 1.00 | 58.06 | B | N |
| ATOM | 2763 | C | ARG | B | 134 | 67.845 | -15.919 | -13.349 | 1.00 | 49.72 | B | C |
| ATOM | 2764 | O | ARG | B | 134 | 67.590 | -16.967 | -13.964 | 1.00 | 50.17 | B | O |
| ATOM | 2765 | N | GLU | B | 135 | 68.864 | -15.116 | -13.650 | 1.00 | 50.18 | B | N |
| ATOM | 2766 | CA | GLU | B | 135 | 69.797 | -15.421 | -14.729 | 1.00 | 50.10 | B | C |
| ATOM | 2767 | CB | GLU | B | 135 | 70.588 | -14.172 | -15.126 | 1.00 | 50.99 | B | C |
| ATOM | 2768 | CG | GLU | B | 135 | 69.747 | -13.098 | -15.787 | 1.00 | 52.01 | B | C |
| ATOM | 2769 | CD | GLU | B | 135 | 69.216 | -13.522 | -17.139 | 1.00 | 52.98 | B | C |

FIGURE 9 - 42

SUBSTITUTE SHEET (RULE 26)

52/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|---------|------|-------|---|---|
| ATOM | 2770 | OE1 | GLU | B | 135 | 69.008 | -14.750 | -17.340 | 1.00 | 53.03 | B | O |
| ATOM | 2771 | OE2 | GLU | B | 135 | 68.994 | -12.625 | -17.995 | 1.00 | 52.29 | B | O |
| ATOM | 2772 | C | GLU | B | 135 | 70.764 | -16.496 | -14.287 | 1.00 | 49.45 | B | C |
| ATOM | 2773 | O | GLU | B | 135 | 71.096 | -17.389 | -15.063 | 1.00 | 49.99 | B | O |
| ATOM | 2774 | N | ALA | B | 136 | 71.223 | -16.409 | -13.041 | 1.00 | 48.59 | B | N |
| ATOM | 2775 | CA | ALA | B | 136 | 72.170 | -17.388 | -12.519 | 1.00 | 47.60 | B | C |
| ATOM | 2776 | CB | ALA | B | 136 | 72.982 | -16.779 | -11.378 | 1.00 | 48.53 | B | C |
| ATOM | 2777 | C | ALA | B | 136 | 71.448 | -18.638 | -12.040 | 1.00 | 47.54 | B | C |
| ATOM | 2778 | O | ALA | B | 136 | 72.033 | -19.490 | -11.369 | 1.00 | 47.42 | B | O |
| ATOM | 2779 | N | GLY | B | 137 | 70.172 | -18.740 | -12.400 | 1.00 | 48.02 | B | N |
| ATOM | 2780 | CA | GLY | B | 137 | 69.366 | -19.886 | -12.016 | 1.00 | 47.64 | B | C |
| ATOM | 2781 | C | GLY | B | 137 | 69.156 | -20.029 | -10.521 | 1.00 | 47.61 | B | C |
| ATOM | 2782 | O | GLY | B | 137 | 69.015 | -21.150 | -10.025 | 1.00 | 47.78 | B | O |
| ATOM | 2783 | N | LYS | B | 138 | 69.127 | -18.912 | -9.797 | 1.00 | 46.70 | B | N |
| ATOM | 2784 | CA | LYS | B | 138 | 68.944 | -18.957 | -8.348 | 1.00 | 46.65 | B | C |
| ATOM | 2785 | CB | LYS | B | 138 | 69.907 | -17.977 | -7.659 | 1.00 | 46.87 | B | C |
| ATOM | 2786 | CG | LYS | B | 138 | 71.042 | -18.660 | -6.872 | 1.00 | 48.78 | B | C |
| ATOM | 2787 | CD | LYS | B | 138 | 72.280 | -18.916 | -7.729 | 1.00 | 49.16 | B | C |
| ATOM | 2788 | CE | LYS | B | 138 | 73.061 | -17.634 | -8.025 | 1.00 | 49.49 | B | C |
| ATOM | 2789 | NZ | LYS | B | 138 | 73.808 | -17.101 | -6.841 | 1.00 | 47.69 | B | N |
| ATOM | 2790 | C | LYS | B | 138 | 67.511 | -18.683 | -7.867 | 1.00 | 45.56 | B | C |
| ATOM | 2791 | O | LYS | B | 138 | 67.253 | -18.654 | -6.661 | 1.00 | 45.13 | B | O |
| ATOM | 2792 | N | THR | B | 139 | 66.591 | -18.486 | -8.805 | 1.00 | 44.29 | B | N |
| ATOM | 2793 | CA | THR | B | 139 | 65.188 | -18.222 | -8.482 | 1.00 | 43.92 | B | C |
| ATOM | 2794 | CB | THR | B | 139 | 64.263 | -18.545 | -9.678 | 1.00 | 43.58 | B | C |
| ATOM | 2795 | OG1 | THR | B | 139 | 64.092 | -17.376 | -10.479 | 1.00 | 42.78 | B | O |
| ATOM | 2796 | CG2 | THR | B | 139 | 62.898 | -19.016 | -9.195 | 1.00 | 43.55 | B | C |
| ATOM | 2797 | C | THR | B | 139 | 64.666 | -19.006 | -7.282 | 1.00 | 43.14 | B | C |
| ATOM | 2798 | O | THR | B | 139 | 64.202 | -18.421 | -6.304 | 1.00 | 42.23 | B | O |
| ATOM | 2799 | N | LEU | B | 140 | 64.739 | -20.331 | -7.371 | 1.00 | 42.83 | B | N |
| ATOM | 2800 | CA | LEU | B | 140 | 64.248 | -21.203 | -6.312 | 1.00 | 43.20 | B | C |
| ATOM | 2801 | CB | LEU | B | 140 | 64.098 | -22.635 | -6.840 | 1.00 | 43.05 | B | C |
| ATOM | 2802 | CG | LEU | B | 140 | 63.054 | -22.858 | -7.948 | 1.00 | 44.80 | B | C |
| ATOM | 2803 | CD1 | LEU | B | 140 | 63.216 | -24.262 | -8.504 | 1.00 | 44.05 | B | C |
| ATOM | 2804 | CD2 | LEU | B | 140 | 61.624 | -22.648 | -7.404 | 1.00 | 42.29 | B | C |
| ATOM | 2805 | C | LEU | B | 140 | 65.130 | -21.202 | -5.076 | 1.00 | 42.63 | B | C |
| ATOM | 2806 | O | LEU | B | 140 | 64.679 | -21.569 | -3.991 | 1.00 | 42.72 | B | O |
| ATOM | 2807 | N | GLU | B | 141 | 66.384 | -20.790 | -5.240 | 1.00 | 41.84 | B | N |
| ATOM | 2808 | CA | GLU | B | 141 | 67.328 | -20.751 | -4.129 | 1.00 | 41.43 | B | C |
| ATOM | 2809 | CB | GLU | B | 141 | 68.766 | -20.772 | -4.665 | 1.00 | 43.70 | B | C |
| ATOM | 2810 | CG | GLU | B | 141 | 69.846 | -20.775 | -3.591 | 1.00 | 46.88 | B | C |
| ATOM | 2811 | CD | GLU | B | 141 | 71.248 | -20.980 | -4.172 | 1.00 | 50.70 | B | C |
| ATOM | 2812 | OE1 | GLU | B | 141 | 71.445 | -21.986 | -4.904 | 1.00 | 50.88 | B | O |
| ATOM | 2813 | OE2 | GLU | B | 141 | 72.150 | -20.144 | -3.893 | 1.00 | 52.00 | B | O |
| ATOM | 2814 | C | GLU | B | 141 | 67.111 | -19.496 | -3.288 | 1.00 | 39.58 | B | C |
| ATOM | 2815 | O | GLU | B | 141 | 67.223 | -19.521 | -2.059 | 1.00 | 39.45 | B | O |
| ATOM | 2816 | N | VAL | B | 142 | 66.804 | -18.394 | -3.957 | 1.00 | 37.70 | B | N |
| ATOM | 2817 | CA | VAL | B | 142 | 66.581 | -17.138 | -3.256 | 1.00 | 36.79 | B | C |
| ATOM | 2818 | CB | VAL | B | 142 | 66.536 | -15.974 | -4.260 | 1.00 | 37.75 | B | C |
| ATOM | 2819 | CG1 | VAL | B | 142 | 66.322 | -14.650 | -3.528 | 1.00 | 37.79 | B | C |
| ATOM | 2820 | CG2 | VAL | B | 142 | 67.849 | -15.938 | -5.037 | 1.00 | 36.29 | B | C |
| ATOM | 2821 | C | VAL | B | 142 | 65.283 | -17.216 | -2.457 | 1.00 | 35.74 | B | C |
| ATOM | 2822 | O | VAL | B | 142 | 65.277 | -17.019 | -1.235 | 1.00 | 35.29 | B | O |
| ATOM | 2823 | N | VAL | B | 143 | 64.185 | -17.513 | -3.142 | 1.00 | 33.65 | B | N |
| ATOM | 2824 | CA | VAL | B | 143 | 62.895 | -17.639 | -2.473 | 1.00 | 33.11 | B | C |
| ATOM | 2825 | CB | VAL | B | 143 | 61.816 | -18.215 | -3.438 | 1.00 | 31.68 | B | C |
| ATOM | 2826 | CG1 | VAL | B | 143 | 60.549 | -18.563 | -2.671 | 1.00 | 29.65 | B | C |
| ATOM | 2827 | CG2 | VAL | B | 143 | 61.497 | -17.198 | -4.525 | 1.00 | 30.74 | B | C |
| ATOM | 2828 | C | VAL | B | 143 | 63.027 | -18.557 | -1.261 | 1.00 | 33.33 | B | C |
| ATOM | 2829 | O | VAL | B | 143 | 62.555 | -18.236 | -0.165 | 1.00 | 34.17 | B | O |
| ATOM | 2830 | N | ALA | B | 144 | 63.669 | -19.704 | -1.462 | 1.00 | 32.42 | B | N |
| ATOM | 2831 | CA | ALA | B | 144 | 63.853 | -20.666 | -0.387 | 1.00 | 32.54 | B | C |
| ATOM | 2832 | CB | ALA | B | 144 | 64.577 | -21.910 | -0.915 | 1.00 | 33.65 | B | C |
| ATOM | 2833 | C | ALA | B | 144 | 64.623 | -20.061 | 0.781 | 1.00 | 31.33 | B | C |
| ATOM | 2834 | O | ALA | B | 144 | 64.309 | -20.329 | 1.936 | 1.00 | 30.69 | B | O |
| ATOM | 2835 | N | ARG | B | 145 | 65.637 | -19.252 | 0.485 | 1.00 | 31.02 | B | N |

FIGURE 9 - 43

SUBSTITUTE SHEET (RULE 26)

53/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|--------|------|-------|---|---|
| ATOM | 2836 | CA | ARG | B | 145 | 66.421 | -18.615 | 1.545 | 1.00 | 31.96 | B | C |
| ATOM | 2837 | CB | ARG | B | 145 | 67.652 | -17.893 | 0.978 | 1.00 | 32.01 | B | C |
| ATOM | 2838 | CG | ARG | B | 145 | 68.436 | -17.157 | 2.059 | 1.00 | 32.50 | B | C |
| ATOM | 2839 | CD | ARG | B | 145 | 69.333 | -16.074 | 1.489 | 1.00 | 34.81 | B | C |
| ATOM | 2840 | NE | ARG | B | 145 | 68.718 | -14.745 | 1.553 | 1.00 | 34.65 | B | N |
| ATOM | 2841 | CZ | ARG | B | 145 | 68.463 | -13.989 | 0.492 | 1.00 | 36.06 | B | N |
| ATOM | 2842 | NH1 | ARG | B | 145 | 68.765 | -14.434 | -0.720 | 1.00 | 36.01 | B | N |
| ATOM | 2843 | NH2 | ARG | B | 145 | 67.925 | -12.779 | 0.644 | 1.00 | 38.36 | B | N |
| ATOM | 2844 | C | ARG | B | 145 | 65.586 | -17.582 | 2.305 | 1.00 | 31.36 | B | C |
| ATOM | 2845 | O | ARG | B | 145 | 65.607 | -17.537 | 3.532 | 1.00 | 31.43 | B | O |
| ATOM | 2846 | N | GLN | B | 146 | 64.869 | -16.742 | 1.567 | 1.00 | 31.80 | B | N |
| ATOM | 2847 | CA | GLN | B | 146 | 64.056 | -15.706 | 2.190 | 1.00 | 31.49 | B | C |
| ATOM | 2848 | CB | GLN | B | 146 | 63.451 | -14.789 | 1.122 | 1.00 | 32.32 | B | C |
| ATOM | 2849 | CG | GLN | B | 146 | 64.490 | -14.024 | 0.291 | 1.00 | 33.24 | B | C |
| ATOM | 2850 | CD | GLN | B | 146 | 63.870 | -13.245 | -0.857 | 1.00 | 33.89 | B | C |
| ATOM | 2851 | OE1 | GLN | B | 146 | 63.141 | -13.807 | -1.679 | 1.00 | 36.66 | B | O |
| ATOM | 2852 | NE2 | GLN | B | 146 | 64.164 | -11.947 | -0.929 | 1.00 | 32.77 | B | N |
| ATOM | 2853 | C | GLN | B | 146 | 62.957 | -16.354 | 3.009 | 1.00 | 32.42 | B | C |
| ATOM | 2854 | O | GLN | B | 146 | 62.724 | -15.984 | 4.158 | 1.00 | 32.36 | B | O |
| ATOM | 2855 | N | LEU | B | 147 | 62.283 | -17.332 | 2.415 | 1.00 | 32.71 | B | N |
| ATOM | 2856 | CA | LEU | B | 147 | 61.209 | -18.021 | 3.111 | 1.00 | 33.14 | B | C |
| ATOM | 2857 | CB | LEU | B | 147 | 60.482 | -18.958 | 2.148 | 1.00 | 33.72 | B | C |
| ATOM | 2858 | CG | LEU | B | 147 | 58.968 | -19.089 | 2.309 | 1.00 | 34.18 | B | C |
| ATOM | 2859 | CD1 | LEU | B | 147 | 58.451 | -20.102 | 1.309 | 1.00 | 35.74 | B | C |
| ATOM | 2860 | CD2 | LEU | B | 147 | 58.619 | -19.517 | 3.706 | 1.00 | 36.43 | B | C |
| ATOM | 2861 | C | LEU | B | 147 | 61.802 | -18.821 | 4.267 | 1.00 | 33.58 | B | C |
| ATOM | 2862 | O | LEU | B | 147 | 61.248 | -18.850 | 5.371 | 1.00 | 32.44 | B | O |
| ATOM | 2863 | N | GLY | B | 148 | 62.939 | -19.462 | 4.007 | 1.00 | 34.35 | B | N |
| ATOM | 2864 | CA | GLY | B | 148 | 63.600 | -20.264 | 5.025 | 1.00 | 33.95 | B | C |
| ATOM | 2865 | C | GLY | B | 148 | 64.010 | -19.498 | 6.271 | 1.00 | 33.58 | B | C |
| ATOM | 2866 | O | GLY | B | 148 | 63.808 | -19.973 | 7.385 | 1.00 | 33.55 | B | O |
| ATOM | 2867 | N | SER | B | 149 | 64.587 | -18.313 | 6.091 | 1.00 | 33.86 | B | N |
| ATOM | 2868 | CA | SER | B | 149 | 65.018 | -17.505 | 7.228 | 1.00 | 34.39 | B | C |
| ATOM | 2869 | CB | SER | B | 149 | 65.411 | -16.102 | 6.774 | 1.00 | 33.76 | B | C |
| ATOM | 2870 | OG | SER | B | 149 | 66.382 | -16.158 | 5.748 | 1.00 | 37.90 | B | O |
| ATOM | 2871 | C | SER | B | 149 | 63.887 | -17.391 | 8.233 | 1.00 | 33.93 | B | C |
| ATOM | 2872 | O | SER | B | 149 | 64.063 | -17.651 | 9.413 | 1.00 | 35.67 | B | O |
| ATOM | 2873 | N | VAL | B | 150 | 62.714 | -17.003 | 7.759 | 1.00 | 33.51 | B | N |
| ATOM | 2874 | CA | VAL | B | 150 | 61.581 | -16.859 | 8.652 | 1.00 | 33.52 | B | C |
| ATOM | 2875 | CB | VAL | B | 150 | 60.380 | -16.249 | 7.900 | 1.00 | 31.38 | B | C |
| ATOM | 2876 | CG1 | VAL | B | 150 | 59.171 | -16.135 | 8.814 | 1.00 | 29.87 | B | C |
| ATOM | 2877 | CG2 | VAL | B | 150 | 60.774 | -14.878 | 7.377 | 1.00 | 29.06 | B | C |
| ATOM | 2878 | C | VAL | B | 150 | 61.227 | -18.204 | 9.288 | 1.00 | 33.67 | B | C |
| ATOM | 2879 | O | VAL | B | 150 | 61.143 | -18.299 | 10.505 | 1.00 | 33.91 | B | O |
| ATOM | 2880 | N | ILE | B | 151 | 61.038 | -19.241 | 8.472 | 1.00 | 35.00 | B | N |
| ATOM | 2881 | CA | ILE | B | 151 | 60.715 | -20.570 | 8.994 | 1.00 | 34.83 | B | C |
| ATOM | 2882 | CB | ILE | B | 151 | 60.713 | -21.645 | 7.874 | 1.00 | 34.84 | B | C |
| ATOM | 2883 | CG2 | ILE | B | 151 | 60.610 | -23.053 | 8.489 | 1.00 | 32.94 | B | C |
| ATOM | 2884 | CG1 | ILE | B | 151 | 59.540 | -21.405 | 6.923 | 1.00 | 34.82 | B | C |
| ATOM | 2885 | CD1 | ILE | B | 151 | 59.536 | -22.306 | 5.717 | 1.00 | 33.61 | B | C |
| ATOM | 2886 | C | ILE | B | 151 | 61.731 | -20.993 | 10.048 | 1.00 | 35.84 | B | C |
| ATOM | 2887 | O | ILE | B | 151 | 61.364 | -21.412 | 11.140 | 1.00 | 35.24 | B | O |
| ATOM | 2888 | N | ASP | B | 152 | 63.013 | -20.876 | 9.718 | 1.00 | 37.41 | B | N |
| ATOM | 2889 | CA | ASP | B | 152 | 64.059 | -21.275 | 10.650 | 1.00 | 39.98 | B | C |
| ATOM | 2890 | CB | ASP | B | 152 | 65.435 | -21.250 | 9.961 | 1.00 | 40.02 | B | C |
| ATOM | 2891 | CG | ASP | B | 152 | 65.618 | -22.385 | 8.936 | 1.00 | 41.78 | B | C |
| ATOM | 2892 | OD1 | ASP | B | 152 | 64.893 | -23.411 | 9.009 | 1.00 | 41.56 | B | O |
| ATOM | 2893 | OD2 | ASP | B | 152 | 66.511 | -22.252 | 8.063 | 1.00 | 42.08 | B | O |
| ATOM | 2894 | C | ASP | B | 152 | 64.096 | -20.402 | 11.908 | 1.00 | 41.19 | B | C |
| ATOM | 2895 | O | ASP | B | 152 | 64.568 | -20.841 | 12.960 | 1.00 | 42.08 | B | O |
| ATOM | 2896 | N | GLU | B | 153 | 63.592 | -19.173 | 11.809 | 1.00 | 41.50 | B | N |
| ATOM | 2897 | CA | GLU | B | 153 | 63.596 | -18.256 | 12.954 | 1.00 | 40.89 | B | C |
| ATOM | 2898 | CB | GLU | B | 153 | 63.871 | -16.828 | 12.475 | 1.00 | 42.30 | B | C |
| ATOM | 2899 | CG | GLU | B | 153 | 65.294 | -16.337 | 12.724 | 1.00 | 45.45 | B | C |
| ATOM | 2900 | CD | GLU | B | 153 | 65.494 | -15.831 | 14.147 | 1.00 | 47.37 | B | C |
| ATOM | 2901 | OE1 | GLU | B | 153 | 65.446 | -16.655 | 15.091 | 1.00 | 46.76 | B | O |

FIGURE 9 - 44

SUBSTITUTE SHEET (RULE 26)

54/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|--------|------|-------|---|---|
| ATOM | 2902 | OE2 | GLU | B | 153 | 65.685 | -14.601 | 14.322 | 1.00 | 46.66 | B | O |
| ATOM | 2903 | C | GLU | B | 153 | 62.304 | -18.285 | 13.763 | 1.00 | 39.67 | B | C |
| ATOM | 2904 | O | GLU | B | 153 | 62.337 | -18.293 | 14.993 | 1.00 | 39.87 | B | O |
| ATOM | 2905 | N | LEU | B | 154 | 61.166 | -18.316 | 13.079 | 1.00 | 38.28 | B | N |
| ATOM | 2906 | CA | LEU | B | 154 | 59.874 | -18.335 | 13.766 | 1.00 | 37.35 | B | C |
| ATOM | 2907 | CB | LEU | B | 154 | 58.958 | -17.287 | 13.144 | 1.00 | 38.56 | B | C |
| ATOM | 2908 | CG | LEU | B | 154 | 59.520 | -15.869 | 13.129 | 1.00 | 38.92 | B | C |
| ATOM | 2909 | CD1 | LEU | B | 154 | 58.747 | -15.043 | 12.138 | 1.00 | 38.21 | B | C |
| ATOM | 2910 | CD2 | LEU | B | 154 | 59.446 | -15.262 | 14.525 | 1.00 | 40.92 | B | C |
| ATOM | 2911 | C | LEU | B | 154 | 59.154 | -19.684 | 13.772 | 1.00 | 36.33 | B | C |
| ATOM | 2912 | O | LEU | B | 154 | 58.220 | -19.890 | 14.550 | 1.00 | 34.85 | B | O |
| ATOM | 2913 | N | GLY | B | 155 | 59.579 | -20.606 | 12.914 | 1.00 | 37.04 | B | N |
| ATOM | 2914 | CA | GLY | B | 155 | 58.912 | -21.902 | 12.860 | 1.00 | 36.15 | B | C |
| ATOM | 2915 | C | GLY | B | 155 | 57.728 | -21.853 | 11.906 | 1.00 | 36.68 | B | C |
| ATOM | 2916 | O | GLY | B | 155 | 56.984 | -20.865 | 11.868 | 1.00 | 36.28 | B | O |
| ATOM | 2917 | N | VAL | B | 156 | 57.545 | -22.920 | 11.133 | 1.00 | 35.24 | B | N |
| ATOM | 2918 | CA | VAL | B | 156 | 56.472 | -22.986 | 10.149 | 1.00 | 33.52 | B | C |
| ATOM | 2919 | CB | VAL | B | 156 | 56.421 | -24.372 | 9.508 | 1.00 | 34.06 | B | C |
| ATOM | 2920 | CG1 | VAL | B | 156 | 55.544 | -24.342 | 8.254 | 1.00 | 29.95 | B | C |
| ATOM | 2921 | CG2 | VAL | B | 156 | 57.829 | -24.813 | 9.199 | 1.00 | 33.87 | B | C |
| ATOM | 2922 | C | VAL | B | 156 | 55.097 | -22.672 | 10.705 | 1.00 | 32.71 | B | C |
| ATOM | 2923 | O | VAL | B | 156 | 54.213 | -22.219 | 9.979 | 1.00 | 31.42 | B | O |
| ATOM | 2924 | N | GLY | B | 157 | 54.912 | -22.926 | 11.992 | 1.00 | 33.45 | B | N |
| ATOM | 2925 | CA | GLY | B | 157 | 53.626 | -22.659 | 12.602 | 1.00 | 33.58 | B | C |
| ATOM | 2926 | C | GLY | B | 157 | 53.209 | -21.206 | 12.448 | 1.00 | 33.84 | B | C |
| ATOM | 2927 | O | GLY | B | 157 | 52.031 | -20.915 | 12.259 | 1.00 | 34.59 | B | O |
| ATOM | 2928 | N | ALA | B | 158 | 54.172 | -20.291 | 12.510 | 1.00 | 33.02 | B | N |
| ATOM | 2929 | CA | ALA | B | 158 | 53.866 | -18.863 | 12.401 | 1.00 | 32.72 | B | C |
| ATOM | 2930 | CB | ALA | B | 158 | 55.148 | -18.031 | 12.551 | 1.00 | 31.21 | B | C |
| ATOM | 2931 | C | ALA | B | 158 | 53.170 | -18.504 | 11.095 | 1.00 | 32.30 | B | C |
| ATOM | 2932 | O | ALA | B | 158 | 52.512 | -17.470 | 11.009 | 1.00 | 32.08 | B | O |
| ATOM | 2933 | N | PHE | B | 159 | 53.319 | -19.352 | 10.079 | 1.00 | 31.55 | B | N |
| ATOM | 2934 | CA | PHE | B | 159 | 52.700 | -19.090 | 8.785 | 1.00 | 30.15 | B | C |
| ATOM | 2935 | CB | PHE | B | 159 | 53.396 | -19.912 | 7.698 | 1.00 | 28.09 | B | C |
| ATOM | 2936 | CG | PHE | B | 159 | 54.661 | -19.275 | 7.183 | 1.00 | 28.95 | B | C |
| ATOM | 2937 | CD1 | PHE | B | 159 | 54.608 | -18.273 | 6.214 | 1.00 | 29.10 | B | C |
| ATOM | 2938 | CD2 | PHE | B | 159 | 55.901 | -19.638 | 7.695 | 1.00 | 28.34 | B | C |
| ATOM | 2939 | CE1 | PHE | B | 159 | 55.770 | -17.641 | 5.764 | 1.00 | 28.03 | B | C |
| ATOM | 2940 | CE2 | PHE | B | 159 | 57.071 | -19.009 | 7.251 | 1.00 | 30.31 | B | C |
| ATOM | 2941 | CZ | PHE | B | 159 | 57.004 | -18.010 | 6.286 | 1.00 | 28.14 | B | C |
| ATOM | 2942 | C | PHE | B | 159 | 51.187 | -19.310 | 8.759 | 1.00 | 29.15 | B | C |
| ATOM | 2943 | O | PHE | B | 159 | 50.516 | -18.900 | 7.816 | 1.00 | 29.06 | B | O |
| ATOM | 2944 | N | ALA | B | 160 | 50.645 | -19.935 | 9.797 | 1.00 | 28.36 | B | N |
| ATOM | 2945 | CA | ALA | B | 160 | 49.199 | -20.155 | 9.869 | 1.00 | 29.28 | B | C |
| ATOM | 2946 | CB | ALA | B | 160 | 48.850 | -21.089 | 11.034 | 1.00 | 29.18 | B | C |
| ATOM | 2947 | C | ALA | B | 160 | 48.518 | -18.806 | 10.062 | 1.00 | 28.57 | B | C |
| ATOM | 2948 | O | ALA | B | 160 | 47.329 | -18.658 | 9.807 | 1.00 | 29.09 | B | O |
| ATOM | 2949 | N | ARG | B | 161 | 49.283 | -17.833 | 10.546 | 1.00 | 29.03 | B | N |
| ATOM | 2950 | CA | ARG | B | 161 | 48.774 | -16.483 | 10.760 | 1.00 | 30.38 | B | C |
| ATOM | 2951 | CB | ARG | B | 161 | 48.880 | -16.097 | 12.233 | 1.00 | 33.51 | B | C |
| ATOM | 2952 | CG | ARG | B | 161 | 48.003 | -16.920 | 13.161 | 1.00 | 38.47 | B | C |
| ATOM | 2953 | CD | ARG | B | 161 | 48.273 | -16.546 | 14.617 | 1.00 | 42.43 | B | C |
| ATOM | 2954 | NE | ARG | B | 161 | 47.547 | -17.410 | 15.556 | 1.00 | 48.87 | B | N |
| ATOM | 2955 | CZ | ARG | B | 161 | 47.620 | -18.744 | 15.577 | 1.00 | 50.97 | B | C |
| ATOM | 2956 | NH1 | ARG | B | 161 | 48.389 | -19.390 | 14.706 | 1.00 | 51.93 | B | N |
| ATOM | 2957 | NH2 | ARG | B | 161 | 46.928 | -19.436 | 16.480 | 1.00 | 53.25 | B | N |
| ATOM | 2958 | C | ARG | B | 161 | 49.537 | -15.463 | 9.908 | 1.00 | 29.14 | B | C |
| ATOM | 2959 | O | ARG | B | 161 | 49.798 | -14.337 | 10.340 | 1.00 | 30.76 | B | O |
| ATOM | 2960 | N | ALA | B | 162 | 49.903 | -15.854 | 8.697 | 1.00 | 26.93 | B | N |
| ATOM | 2961 | CA | ALA | B | 162 | 50.609 | -14.940 | 7.819 | 1.00 | 25.10 | B | C |
| ATOM | 2962 | CB | ALA | B | 162 | 52.103 | -15.147 | 7.940 | 1.00 | 23.97 | B | C |
| ATOM | 2963 | C | ALA | B | 162 | 50.162 | -15.157 | 6.385 | 1.00 | 25.40 | B | C |
| ATOM | 2964 | O | ALA | B | 162 | 49.250 | -15.948 | 6.123 | 1.00 | 26.64 | B | O |
| ATOM | 2965 | N | VAL | B | 163 | 50.807 | -14.434 | 5.472 | 1.00 | 24.49 | B | N |
| ATOM | 2966 | CA | VAL | B | 163 | 50.548 | -14.507 | 4.038 | 1.00 | 22.85 | B | C |
| ATOM | 2967 | CB | VAL | B | 163 | 49.520 | -13.426 | 3.611 | 1.00 | 21.91 | B | C |

FIGURE 9 - 45

55/76

| | | | | | | | | | |
|------|------|---------------|--------|---------|---------|------|-------|---|---|
| ATOM | 2968 | CG1 VAL B 163 | 49.464 | -13.308 | 2.103 | 1.00 | 19.95 | B | C |
| ATOM | 2969 | CG2 VAL B 163 | 48.137 | -13.784 | 4.163 | 1.00 | 21.94 | B | C |
| ATOM | 2970 | C VAL B 163 | 51.889 | -14.257 | 3.334 | 1.00 | 23.41 | B | C |
| ATOM | 2971 | O VAL B 163 | 52.754 | -13.564 | 3.879 | 1.00 | 23.24 | B | O |
| ATOM | 2972 | N VAL B 164 | 52.078 | -14.847 | 2.154 | 1.00 | 24.25 | B | N |
| ATOM | 2973 | CA VAL B 164 | 53.313 | -14.640 | 1.391 | 1.00 | 23.50 | B | C |
| ATOM | 2974 | CB VAL B 164 | 54.103 | -15.961 | 1.184 | 1.00 | 24.53 | B | C |
| ATOM | 2975 | CG1 VAL B 164 | 55.356 | -15.708 | 0.339 | 1.00 | 23.22 | B | C |
| ATOM | 2976 | CG2 VAL B 164 | 54.495 | -16.545 | 2.518 | 1.00 | 21.44 | B | C |
| ATOM | 2977 | C VAL B 164 | 52.984 | -14.043 | 0.027 | 1.00 | 25.86 | B | C |
| ATOM | 2978 | O VAL B 164 | 51.912 | -14.296 | -0.528 | 1.00 | 27.41 | B | O |
| ATOM | 2979 | N ALA B 165 | 53.903 | -13.236 | -0.501 | 1.00 | 26.27 | B | N |
| ATOM | 2980 | CA ALA B 165 | 53.727 | -12.605 | -1.799 | 1.00 | 26.39 | B | C |
| ATOM | 2981 | CB ALA B 165 | 53.356 | -11.125 | -1.626 | 1.00 | 26.44 | B | C |
| ATOM | 2982 | C ALA B 165 | 55.011 | -12.737 | -2.616 | 1.00 | 27.67 | B | C |
| ATOM | 2983 | O ALA B 165 | 56.089 | -12.301 | -2.194 | 1.00 | 26.72 | B | O |
| ATOM | 2984 | N TYR B 166 | 54.894 | -13.360 | -3.783 | 1.00 | 27.21 | B | N |
| ATOM | 2985 | CA TYR B 166 | 56.038 | -13.540 | -4.652 | 1.00 | 25.91 | B | C |
| ATOM | 2986 | CB TYR B 166 | 55.917 | -14.847 | -5.445 | 1.00 | 25.86 | B | C |
| ATOM | 2987 | CG TYR B 166 | 56.958 | -15.001 | -6.542 | 1.00 | 26.86 | B | C |
| ATOM | 2988 | CD1 TYR B 166 | 58.326 | -14.907 | -6.260 | 1.00 | 27.39 | B | C |
| ATOM | 2989 | CE1 TYR B 166 | 59.283 | -14.982 | -7.283 | 1.00 | 27.38 | B | C |
| ATOM | 2990 | CD2 TYR B 166 | 56.574 | -15.188 | -7.875 | 1.00 | 26.41 | B | C |
| ATOM | 2991 | CE2 TYR B 166 | 57.524 | -15.267 | -8.900 | 1.00 | 27.40 | B | C |
| ATOM | 2992 | CZ TYR B 166 | 58.872 | -15.158 | -8.600 | 1.00 | 28.42 | B | C |
| ATOM | 2993 | OH TYR B 166 | 59.799 | -15.186 | -9.626 | 1.00 | 29.66 | B | O |
| ATOM | 2994 | C TYR B 166 | 56.149 | -12.362 | -5.606 | 1.00 | 26.50 | B | C |
| ATOM | 2995 | O TYR B 166 | 55.275 | -12.136 | -6.443 | 1.00 | 26.34 | B | O |
| ATOM | 2996 | N GLU B 167 | 57.221 | -11.597 | -5.453 | 1.00 | 27.03 | B | N |
| ATOM | 2997 | CA GLU B 167 | 57.472 | -10.450 | -6.312 | 1.00 | 27.72 | B | C |
| ATOM | 2998 | CB GLU B 167 | 58.069 | -9.274 | -5.540 | 1.00 | 28.49 | B | C |
| ATOM | 2999 | CG GLU B 167 | 57.234 | -8.638 | -4.463 | 1.00 | 30.49 | B | C |
| ATOM | 3000 | CD GLU B 167 | 57.965 | -7.429 | -3.881 | 1.00 | 31.03 | B | C |
| ATOM | 3001 | OE1 GLU B 167 | 58.238 | -6.472 | -4.637 | 1.00 | 30.63 | B | O |
| ATOM | 3002 | OE2 GLU B 167 | 58.284 | -7.445 | -2.677 | 1.00 | 30.55 | B | O |
| ATOM | 3003 | C GLU B 167 | 58.521 | -10.858 | -7.323 | 1.00 | 28.43 | B | C |
| ATOM | 3004 | O GLU B 167 | 59.666 | -11.126 | -6.953 | 1.00 | 27.47 | B | O |
| ATOM | 3005 | N PRO B 168 | 58.155 | -10.925 | -8.607 | 1.00 | 29.57 | B | N |
| ATOM | 3006 | CD PRO B 168 | 56.823 | -10.892 | -9.233 | 1.00 | 29.71 | B | C |
| ATOM | 3007 | CA PRO B 168 | 59.190 | -11.305 | -9.568 | 1.00 | 31.96 | B | C |
| ATOM | 3008 | CB PRO B 168 | 58.385 | -11.803 | -10.761 | 1.00 | 31.26 | B | C |
| ATOM | 3009 | CG PRO B 168 | 57.156 | -10.926 | -10.712 | 1.00 | 30.26 | B | C |
| ATOM | 3010 | C PRO B 168 | 59.982 | -10.038 | -9.876 | 1.00 | 33.86 | B | C |
| ATOM | 3011 | O PRO B 168 | 59.731 | -9.375 | -10.879 | 1.00 | 35.18 | B | O |
| ATOM | 3012 | N VAL B 169 | 60.920 | -9.692 | -8.995 | 1.00 | 35.62 | B | N |
| ATOM | 3013 | CA VAL B 169 | 61.747 | -8.493 | -9.167 | 1.00 | 37.76 | B | C |
| ATOM | 3014 | CB VAL B 169 | 62.893 | -8.464 | -8.108 | 1.00 | 37.77 | B | C |
| ATOM | 3015 | CG1 VAL B 169 | 63.881 | -7.360 | -8.425 | 1.00 | 38.52 | B | C |
| ATOM | 3016 | CG2 VAL B 169 | 62.309 | -8.252 | -6.709 | 1.00 | 37.87 | B | C |
| ATOM | 3017 | C VAL B 169 | 62.340 | -8.400 | -10.578 | 1.00 | 39.18 | B | C |
| ATOM | 3018 | O VAL B 169 | 62.525 | -7.309 | -11.118 | 1.00 | 39.32 | B | O |
| ATOM | 3019 | N TRP B 170 | 62.632 | -9.549 | -11.176 | 1.00 | 41.82 | B | N |
| ATOM | 3020 | CA TRP B 170 | 63.194 | -9.588 | -12.522 | 1.00 | 44.07 | B | C |
| ATOM | 3021 | CB TRP B 170 | 63.504 | -11.029 | -12.938 | 1.00 | 44.78 | B | C |
| ATOM | 3022 | CG TRP B 170 | 62.351 | -11.987 | -12.813 | 1.00 | 45.78 | B | C |
| ATOM | 3023 | CD2 TRP B 170 | 61.477 | -12.417 | -13.862 | 1.00 | 47.54 | B | C |
| ATOM | 3024 | CE2 TRP B 170 | 60.564 | -13.338 | -13.296 | 1.00 | 47.68 | B | C |
| ATOM | 3025 | CE3 TRP B 170 | 61.375 | -12.115 | -15.228 | 1.00 | 48.19 | B | C |
| ATOM | 3026 | CD1 TRP B 170 | 61.945 | -12.643 | -11.684 | 1.00 | 46.39 | B | C |
| ATOM | 3027 | NE1 TRP B 170 | 60.873 | -13.460 | -11.967 | 1.00 | 47.25 | B | N |
| ATOM | 3028 | CZ2 TRP B 170 | 59.560 | -13.961 | -14.049 | 1.00 | 49.29 | B | C |
| ATOM | 3029 | CZ3 TRP B 170 | 60.380 | -12.734 | -15.978 | 1.00 | 49.26 | B | C |
| ATOM | 3030 | CH2 TRP B 170 | 59.483 | -13.649 | -15.384 | 1.00 | 49.74 | B | C |
| ATOM | 3031 | C TRP B 170 | 62.293 | -8.964 | -13.576 | 1.00 | 45.28 | B | C |
| ATOM | 3032 | O TRP B 170 | 62.763 | -8.265 | -14.470 | 1.00 | 45.73 | B | O |
| ATOM | 3033 | N ALA B 171 | 60.996 | -9.211 | -13.489 | 1.00 | 47.28 | B | N |

FIGURE 9 - 46

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|---------|------|-------|---|---|
| ATOM | 3034 | CA | ALA | B | 171 | 60.098 | -8.658 | -14.492 | 1.00 | 49.80 | B | C |
| ATOM | 3035 | CB | ALA | B | 171 | 58.903 | -9.568 | -14.669 | 1.00 | 49.37 | B | C |
| ATOM | 3036 | C | ALA | B | 171 | 59.632 | -7.251 | -14.148 | 1.00 | 51.89 | B | C |
| ATOM | 3037 | O | ALA | B | 171 | 58.445 | -6.937 | -14.285 | 1.00 | 53.16 | B | O |
| ATOM | 3038 | N | ILE | B | 172 | 60.555 | -6.400 | -13.701 | 1.00 | 53.42 | B | N |
| ATOM | 3039 | CA | ILE | B | 172 | 60.198 | -5.020 | -13.351 | 1.00 | 55.12 | B | C |
| ATOM | 3040 | CB | ILE | B | 172 | 60.597 | -4.698 | -11.892 | 1.00 | 55.25 | B | C |
| ATOM | 3041 | CG2 | ILE | B | 172 | 60.362 | -3.216 | -11.583 | 1.00 | 54.98 | B | C |
| ATOM | 3042 | CG1 | ILE | B | 172 | 59.771 | -5.561 | -10.932 | 1.00 | 55.57 | B | C |
| ATOM | 3043 | CD1 | ILE | B | 172 | 60.145 | -5.362 | -9.479 | 1.00 | 56.54 | B | C |
| ATOM | 3044 | C | ILE | B | 172 | 60.873 | -4.027 | -14.295 | 1.00 | 56.54 | B | C |
| ATOM | 3045 | O | ILE | B | 172 | 62.071 | -3.751 | -14.169 | 1.00 | 56.11 | B | O |
| ATOM | 3046 | N | GLY | B | 173 | 60.095 | -3.509 | -15.248 | 1.00 | 58.11 | B | N |
| ATOM | 3047 | CA | GLY | B | 173 | 60.606 | -2.553 | -16.214 | 1.00 | 60.26 | B | C |
| ATOM | 3048 | C | GLY | B | 173 | 61.921 | -2.950 | -16.867 | 1.00 | 62.01 | B | C |
| ATOM | 3049 | O | GLY | B | 173 | 62.510 | -2.163 | -17.638 | 1.00 | 62.13 | B | O |
| ATOM | 3050 | N | THR | B | 174 | 62.379 | -4.171 | -16.585 | 1.00 | 62.92 | B | N |
| ATOM | 3051 | CA | THR | B | 174 | 63.637 | -4.663 | -17.148 | 1.00 | 63.93 | B | C |
| ATOM | 3052 | CB | THR | B | 174 | 64.183 | -5.878 | -16.330 | 1.00 | 64.54 | B | C |
| ATOM | 3053 | OG1 | THR | B | 174 | 65.603 | -6.005 | -16.534 | 1.00 | 65.21 | B | O |
| ATOM | 3054 | CG2 | THR | B | 174 | 63.493 | -7.182 | -16.779 | 1.00 | 63.95 | B | C |
| ATOM | 3055 | C | THR | B | 174 | 63.445 | -5.085 | -18.614 | 1.00 | 63.82 | B | C |
| ATOM | 3056 | O | THR | B | 174 | 64.307 | -4.822 | -19.465 | 1.00 | 63.79 | B | O |
| ATOM | 3057 | N | GLY | B | 175 | 62.312 | -5.723 | -18.907 | 1.00 | 63.47 | B | N |
| ATOM | 3058 | CA | GLY | B | 175 | 62.048 | -6.178 | -20.266 | 1.00 | 62.81 | B | C |
| ATOM | 3059 | C | GLY | B | 175 | 61.498 | -7.595 | -20.282 | 1.00 | 62.61 | B | C |
| ATOM | 3060 | O | GLY | B | 175 | 60.755 | -7.980 | -21.199 | 1.00 | 62.39 | B | O |
| ATOM | 3061 | N | LEU | B | 176 | 61.877 | -8.393 | -19.284 | 1.00 | 61.59 | B | N |
| ATOM | 3062 | CA | LEU | B | 176 | 61.371 | -9.764 | -19.185 | 1.00 | 60.50 | B | C |
| ATOM | 3063 | CB | LEU | B | 176 | 62.196 | -10.563 | -18.165 | 1.00 | 60.36 | B | C |
| ATOM | 3064 | CG | LEU | B | 176 | 63.698 | -10.771 | -18.430 | 1.00 | 60.12 | B | C |
| ATOM | 3065 | CD1 | LEU | B | 176 | 64.357 | -11.347 | -17.176 | 1.00 | 59.37 | B | C |
| ATOM | 3066 | CD2 | LEU | B | 176 | 63.905 | -11.711 | -19.626 | 1.00 | 59.19 | B | C |
| ATOM | 3067 | C | LEU | B | 176 | 59.919 | -9.636 | -18.696 | 1.00 | 59.84 | B | C |
| ATOM | 3068 | O | LEU | B | 176 | 59.469 | -8.519 | -18.382 | 1.00 | 60.15 | B | O |
| ATOM | 3069 | N | THR | B | 177 | 59.184 | -10.754 | -18.643 | 1.00 | 58.15 | B | N |
| ATOM | 3070 | CA | THR | B | 177 | 57.794 | -10.738 | -18.162 | 1.00 | 56.05 | B | C |
| ATOM | 3071 | CB | THR | B | 177 | 56.823 | -10.243 | -19.271 | 1.00 | 56.89 | B | C |
| ATOM | 3072 | OG1 | THR | B | 177 | 57.025 | -8.833 | -19.478 | 1.00 | 56.95 | B | O |
| ATOM | 3073 | CG2 | THR | B | 177 | 55.362 | -10.483 | -18.869 | 1.00 | 56.54 | B | C |
| ATOM | 3074 | C | THR | B | 177 | 57.338 | -12.103 | -17.619 | 1.00 | 53.43 | B | C |
| ATOM | 3075 | O | THR | B | 177 | 57.694 | -13.148 | -18.167 | 1.00 | 53.56 | B | O |
| ATOM | 3076 | N | ALA | B | 178 | 56.550 | -12.084 | -16.543 | 1.00 | 50.38 | B | N |
| ATOM | 3077 | CA | ALA | B | 178 | 56.078 | -13.318 | -15.903 | 1.00 | 48.18 | B | C |
| ATOM | 3078 | CB | ALA | B | 178 | 56.136 | -13.168 | -14.390 | 1.00 | 48.52 | B | C |
| ATOM | 3079 | C | ALA | B | 178 | 54.690 | -13.801 | -16.304 | 1.00 | 46.31 | B | C |
| ATOM | 3080 | O | ALA | B | 178 | 53.695 | -13.083 | -16.145 | 1.00 | 46.31 | B | O |
| ATOM | 3081 | N | SER | B | 179 | 54.642 | -15.039 | -16.793 | 1.00 | 43.20 | B | N |
| ATOM | 3082 | CA | SER | B | 179 | 53.412 | -15.690 | -17.227 | 1.00 | 41.17 | B | C |
| ATOM | 3083 | CB | SER | B | 179 | 53.736 | -16.740 | -18.277 | 1.00 | 40.66 | B | C |
| ATOM | 3084 | OG | SER | B | 179 | 54.528 | -17.761 | -17.687 | 1.00 | 39.16 | B | O |
| ATOM | 3085 | C | SER | B | 179 | 52.763 | -16.399 | -16.044 | 1.00 | 40.48 | B | C |
| ATOM | 3086 | O | SER | B | 179 | 53.422 | -16.682 | -15.039 | 1.00 | 39.23 | B | O |
| ATOM | 3087 | N | PRO | B | 180 | 51.461 | -16.708 | -16.153 | 1.00 | 39.69 | B | N |
| ATOM | 3088 | CD | PRO | B | 180 | 50.504 | -16.315 | -17.204 | 1.00 | 39.91 | B | C |
| ATOM | 3089 | CA | PRO | B | 180 | 50.784 | -17.393 | -15.051 | 1.00 | 38.93 | B | C |
| ATOM | 3090 | CB | PRO | B | 180 | 49.321 | -17.469 | -15.521 | 1.00 | 39.22 | B | C |
| ATOM | 3091 | CG | PRO | B | 180 | 49.407 | -17.321 | -17.016 | 1.00 | 40.17 | B | C |
| ATOM | 3092 | C | PRO | B | 180 | 51.396 | -18.758 | -14.750 | 1.00 | 37.09 | B | C |
| ATOM | 3093 | O | PRO | B | 180 | 51.335 | -19.237 | -13.623 | 1.00 | 36.21 | B | O |
| ATOM | 3094 | N | ALA | B | 181 | 52.005 | -19.367 | -15.760 | 1.00 | 35.34 | B | N |
| ATOM | 3095 | CA | ALA | B | 181 | 52.627 | -20.669 | -15.585 | 1.00 | 34.73 | B | C |
| ATOM | 3096 | CB | ALA | B | 181 | 52.934 | -21.287 | -16.947 | 1.00 | 34.77 | B | C |
| ATOM | 3097 | C | ALA | B | 181 | 53.907 | -20.548 | -14.761 | 1.00 | 34.88 | B | C |
| ATOM | 3098 | O | ALA | B | 181 | 54.139 | -21.338 | -13.834 | 1.00 | 35.30 | B | O |
| ATOM | 3099 | N | GLN | B | 182 | 54.735 | -19.560 | -15.102 | 1.00 | 33.69 | B | N |

FIGURE 9 - 47

SUBSTITUTE SHEET (RULE 26)

57/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|---------|------|-------|---|---|
| ATOM | 3100 | CA | GLN | B | 182 | 55.988 | -19.327 | -14.395 | 1.00 | 31.64 | B | C |
| ATOM | 3101 | CB | GLN | B | 182 | 56.782 | -18.224 | -15.090 | 1.00 | 34.31 | B | C |
| ATOM | 3102 | CG | GLN | B | 182 | 57.297 | -18.611 | -16.460 | 1.00 | 36.71 | B | C |
| ATOM | 3103 | CD | GLN | B | 182 | 58.078 | -17.488 | -17.109 | 1.00 | 40.11 | B | C |
| ATOM | 3104 | OE1 | GLN | B | 182 | 57.509 | -16.464 | -17.504 | 1.00 | 42.23 | B | O |
| ATOM | 3105 | NE2 | GLN | B | 182 | 59.393 | -17.664 | -17.211 | 1.00 | 41.43 | B | N |
| ATOM | 3106 | C | GLN | B | 182 | 55.731 | -18.941 | -12.947 | 1.00 | 30.34 | B | C |
| ATOM | 3107 | O | GLN | B | 182 | 56.480 | -19.330 | -12.052 | 1.00 | 29.53 | B | O |
| ATOM | 3108 | N | ALA | B | 183 | 54.675 | -18.168 | -12.716 | 1.00 | 29.37 | B | N |
| ATOM | 3109 | CA | ALA | B | 183 | 54.325 | -17.760 | -11.360 | 1.00 | 29.56 | B | C |
| ATOM | 3110 | CB | ALA | B | 183 | 53.264 | -16.651 | -11.393 | 1.00 | 28.23 | B | C |
| ATOM | 3111 | C | ALA | B | 183 | 53.792 | -18.981 | -10.605 | 1.00 | 29.40 | B | C |
| ATOM | 3112 | O | ALA | B | 183 | 54.158 | -19.224 | -9.457 | 1.00 | 29.35 | B | O |
| ATOM | 3113 | N | GLN | B | 184 | 52.930 | -19.752 | -11.263 | 1.00 | 30.49 | B | N |
| ATOM | 3114 | CA | GLN | B | 184 | 52.359 | -20.948 | -10.656 | 1.00 | 30.41 | B | C |
| ATOM | 3115 | CB | GLN | B | 184 | 51.423 | -21.635 | -11.658 | 1.00 | 29.87 | B | C |
| ATOM | 3116 | CG | GLN | B | 184 | 50.941 | -23.037 | -11.275 | 1.00 | 28.92 | B | C |
| ATOM | 3117 | CD | GLN | B | 184 | 50.270 | -23.109 | -9.911 | 1.00 | 28.37 | B | C |
| ATOM | 3118 | OE1 | GLN | B | 184 | 49.463 | -22.256 | -9.549 | 1.00 | 28.04 | B | O |
| ATOM | 3119 | NE2 | GLN | B | 184 | 50.597 | -24.145 | -9.156 | 1.00 | 26.97 | B | N |
| ATOM | 3120 | C | GLN | B | 184 | 53.470 | -21.896 | -10.193 | 1.00 | 30.23 | B | C |
| ATOM | 3121 | O | GLN | B | 184 | 53.350 | -22.543 | -9.149 | 1.00 | 30.32 | B | O |
| ATOM | 3122 | N | GLU | B | 185 | 54.564 | -21.956 | -10.949 | 1.00 | 30.83 | B | N |
| ATOM | 3123 | CA | GLU | B | 185 | 55.675 | -22.833 | -10.585 | 1.00 | 31.41 | B | C |
| ATOM | 3124 | CB | GLU | B | 185 | 56.746 | -22.878 | -11.673 | 1.00 | 33.38 | B | C |
| ATOM | 3125 | CG | GLU | B | 185 | 56.313 | -23.541 | -12.948 | 1.00 | 38.73 | B | C |
| ATOM | 3126 | CD | GLU | B | 185 | 57.501 | -23.979 | -13.792 | 1.00 | 42.29 | B | C |
| ATOM | 3127 | OE1 | GLU | B | 185 | 58.607 | -23.403 | -13.617 | 1.00 | 43.87 | B | O |
| ATOM | 3128 | OE2 | GLU | B | 185 | 57.325 | -24.890 | -14.635 | 1.00 | 42.49 | B | O |
| ATOM | 3129 | C | GLU | B | 185 | 56.335 | -22.408 | -9.297 | 1.00 | 28.92 | B | C |
| ATOM | 3130 | O | GLU | B | 185 | 56.619 | -23.249 | -8.447 | 1.00 | 27.81 | B | O |
| ATOM | 3131 | N | VAL | B | 186 | 56.601 | -21.109 | -9.173 | 1.00 | 28.30 | B | N |
| ATOM | 3132 | CA | VAL | B | 186 | 57.233 | -20.568 | -7.976 | 1.00 | 28.55 | B | C |
| ATOM | 3133 | CB | VAL | B | 186 | 57.555 | -19.045 | -8.114 | 1.00 | 29.78 | B | C |
| ATOM | 3134 | CG1 | VAL | B | 186 | 58.008 | -18.497 | -6.774 | 1.00 | 28.81 | B | C |
| ATOM | 3135 | CG2 | VAL | B | 186 | 58.648 | -18.813 | -9.149 | 1.00 | 29.17 | B | C |
| ATOM | 3136 | C | VAL | B | 186 | 56.306 | -20.757 | -6.773 | 1.00 | 27.55 | B | C |
| ATOM | 3137 | O | VAL | B | 186 | 56.729 | -21.269 | -5.735 | 1.00 | 27.55 | B | O |
| ATOM | 3138 | N | HIS | B | 187 | 55.048 | -20.341 | -6.909 | 1.00 | 26.98 | B | N |
| ATOM | 3139 | CA | HIS | B | 187 | 54.096 | -20.485 | -5.810 | 1.00 | 27.85 | B | C |
| ATOM | 3140 | CB | HIS | B | 187 | 52.685 | -20.032 | -6.230 | 1.00 | 28.65 | B | C |
| ATOM | 3141 | CG | HIS | B | 187 | 52.523 | -18.546 | -6.332 | 1.00 | 28.35 | B | C |
| ATOM | 3142 | CD2 | HIS | B | 187 | 53.308 | -17.594 | -6.896 | 1.00 | 29.26 | B | C |
| ATOM | 3143 | ND1 | HIS | B | 187 | 51.418 | -17.885 | -5.839 | 1.00 | 28.76 | B | N |
| ATOM | 3144 | CE1 | HIS | B | 187 | 51.529 | -16.591 | -6.095 | 1.00 | 28.46 | B | C |
| ATOM | 3145 | NE2 | HIS | B | 187 | 52.666 | -16.389 | -6.735 | 1.00 | 28.04 | B | N |
| ATOM | 3146 | C | HIS | B | 187 | 54.043 | -21.934 | -5.311 | 1.00 | 28.21 | B | C |
| ATOM | 3147 | O | HIS | B | 187 | 53.917 | -22.174 | -4.109 | 1.00 | 27.87 | B | O |
| ATOM | 3148 | N | ALA | B | 188 | 54.145 | -22.896 | -6.230 | 1.00 | 28.24 | B | N |
| ATOM | 3149 | CA | ALA | B | 188 | 54.110 | -24.310 | -5.843 | 1.00 | 28.15 | B | C |
| ATOM | 3150 | CB | ALA | B | 188 | 53.946 | -25.199 | -7.074 | 1.00 | 29.39 | B | C |
| ATOM | 3151 | C | ALA | B | 188 | 55.372 | -24.701 | -5.075 | 1.00 | 26.86 | B | O |
| ATOM | 3152 | O | ALA | B | 188 | 55.301 | -25.441 | -4.099 | 1.00 | 25.84 | B | N |
| ATOM | 3153 | N | ALA | B | 189 | 56.521 | -24.203 | -5.524 | 1.00 | 25.71 | B | C |
| ATOM | 3154 | CA | ALA | B | 189 | 57.792 | -24.485 | -4.865 | 1.00 | 24.90 | B | C |
| ATOM | 3155 | CB | ALA | B | 189 | 58.946 | -23.830 | -5.637 | 1.00 | 25.61 | B | C |
| ATOM | 3156 | C | ALA | B | 189 | 57.719 | -23.912 | -3.453 | 1.00 | 25.06 | B | O |
| ATOM | 3157 | O | ALA | B | 189 | 58.159 | -24.533 | -2.486 | 1.00 | 23.09 | B | N |
| ATOM | 3158 | N | ILE | B | 190 | 57.161 | -22.711 | -3.343 | 1.00 | 26.46 | B | C |
| ATOM | 3159 | CA | ILE | B | 190 | 57.012 | -22.074 | -2.047 | 1.00 | 26.43 | B | C |
| ATOM | 3160 | CB | ILE | B | 190 | 56.351 | -20.675 | -2.182 | 1.00 | 26.87 | B | C |
| ATOM | 3161 | CG2 | ILE | B | 190 | 55.826 | -20.212 | -0.835 | 1.00 | 27.22 | B | C |
| ATOM | 3162 | CG1 | ILE | B | 190 | 57.357 | -19.673 | -2.758 | 1.00 | 25.60 | B | C |
| ATOM | 3163 | CD1 | ILE | B | 190 | 56.762 | -18.320 | -3.063 | 1.00 | 25.00 | B | C |
| ATOM | 3164 | C | ILE | B | 190 | 56.143 | -22.957 | -1.154 | 1.00 | 26.70 | B | C |
| ATOM | 3165 | O | ILE | B | 190 | 56.548 | -23.329 | -0.049 | 1.00 | 27.43 | B | O |

FIGURE 9 - 48

SUBSTITUTE SHEET (RULE 26)

58/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|--------|------|-------|---|---|
| ATOM | 3166 | N | ARG | B | 191 | 54.956 | -23.308 | -1.642 | 1.00 | 25.64 | B | N |
| ATOM | 3167 | CA | ARG | B | 191 | 54.041 | -24.132 | -0.863 | 1.00 | 26.17 | B | C |
| ATOM | 3168 | CB | ARG | B | 191 | 52.713 | -24.314 | -1.610 | 1.00 | 23.77 | B | C |
| ATOM | 3169 | CG | ARG | B | 191 | 51.527 | -24.471 | -0.682 | 1.00 | 24.78 | B | C |
| ATOM | 3170 | CD | ARG | B | 191 | 50.222 | -24.654 | -1.440 | 1.00 | 25.41 | B | C |
| ATOM | 3171 | NE | ARG | B | 191 | 49.731 | -23.438 | -2.084 | 1.00 | 26.31 | B | N |
| ATOM | 3172 | CZ | ARG | B | 191 | 49.117 | -22.434 | -1.457 | 1.00 | 26.22 | B | C |
| ATOM | 3173 | NH1 | ARG | B | 191 | 48.913 | -22.476 | -0.148 | 1.00 | 22.57 | B | N |
| ATOM | 3174 | NH2 | ARG | B | 191 | 48.669 | -21.400 | -2.155 | 1.00 | 26.06 | B | N |
| ATOM | 3175 | C | ARG | B | 191 | 54.661 | -25.492 | -0.524 | 1.00 | 26.48 | B | C |
| ATOM | 3176 | O | ARG | B | 191 | 54.312 | -26.102 | 0.486 | 1.00 | 27.10 | B | O |
| ATOM | 3177 | N | ALA | B | 192 | 55.592 | -25.953 | -1.357 | 1.00 | 26.70 | B | N |
| ATOM | 3178 | CA | ALA | B | 192 | 56.265 | -27.229 | -1.127 | 1.00 | 27.50 | B | C |
| ATOM | 3179 | CB | ALA | B | 192 | 57.010 | -27.656 | -2.374 | 1.00 | 25.81 | B | C |
| ATOM | 3180 | C | ALA | B | 192 | 57.239 | -27.102 | 0.040 | 1.00 | 29.38 | B | C |
| ATOM | 3181 | O | ALA | B | 192 | 57.342 | -27.998 | 0.890 | 1.00 | 28.55 | B | O |
| ATOM | 3182 | N | GLN | B | 193 | 57.948 | -25.977 | 0.078 | 1.00 | 30.48 | B | N |
| ATOM | 3183 | CA | GLN | B | 193 | 58.916 | -25.726 | 1.133 | 1.00 | 31.50 | B | C |
| ATOM | 3184 | CB | GLN | B | 193 | 59.649 | -24.405 | 0.874 | 1.00 | 33.28 | B | C |
| ATOM | 3185 | CG | GLN | B | 193 | 61.021 | -24.323 | 1.549 | 1.00 | 35.78 | B | C |
| ATOM | 3186 | CD | GLN | B | 193 | 61.643 | -22.937 | 1.481 | 1.00 | 36.05 | B | C |
| ATOM | 3187 | OE1 | GLN | B | 193 | 61.711 | -22.312 | 0.417 | 1.00 | 36.19 | B | O |
| ATOM | 3188 | NE2 | GLN | B | 193 | 62.113 | -22.456 | 2.624 | 1.00 | 35.91 | B | N |
| ATOM | 3189 | C | GLN | B | 193 | 58.180 | -25.678 | 2.472 | 1.00 | 31.70 | B | C |
| ATOM | 3190 | O | GLN | B | 193 | 58.654 | -26.218 | 3.485 | 1.00 | 32.62 | B | O |
| ATOM | 3191 | N | LEU | B | 194 | 57.009 | -25.047 | 2.471 | 1.00 | 30.71 | B | N |
| ATOM | 3192 | CA | LEU | B | 194 | 56.205 | -24.942 | 3.679 | 1.00 | 30.17 | B | C |
| ATOM | 3193 | CB | LEU | B | 194 | 55.020 | -23.992 | 3.451 | 1.00 | 30.51 | B | C |
| ATOM | 3194 | CG | LEU | B | 194 | 55.353 | -22.507 | 3.231 | 1.00 | 31.77 | B | C |
| ATOM | 3195 | CD1 | LEU | B | 194 | 54.078 | -21.720 | 3.021 | 1.00 | 32.18 | B | C |
| ATOM | 3196 | CD2 | LEU | B | 194 | 56.116 | -21.958 | 4.432 | 1.00 | 31.27 | B | C |
| ATOM | 3197 | C | LEU | B | 194 | 55.700 | -26.320 | 4.108 | 1.00 | 29.61 | B | C |
| ATOM | 3198 | O | LEU | B | 194 | 55.689 | -26.647 | 5.293 | 1.00 | 28.17 | B | O |
| ATOM | 3199 | N | ALA | B | 195 | 55.307 | -27.131 | 3.132 | 1.00 | 28.96 | B | N |
| ATOM | 3200 | CA | ALA | B | 195 | 54.790 | -28.467 | 3.401 | 1.00 | 29.13 | B | C |
| ATOM | 3201 | CB | ALA | B | 195 | 54.118 | -29.014 | 2.147 | 1.00 | 27.53 | B | C |
| ATOM | 3202 | C | ALA | B | 195 | 55.857 | -29.451 | 3.900 | 1.00 | 29.18 | B | C |
| ATOM | 3203 | O | ALA | B | 195 | 55.551 | -30.362 | 4.672 | 1.00 | 29.45 | B | O |
| ATOM | 3204 | N | ALA | B | 196 | 57.099 | -29.274 | 3.451 | 1.00 | 29.67 | B | N |
| ATOM | 3205 | CA | ALA | B | 196 | 58.197 | -30.145 | 3.868 | 1.00 | 32.45 | B | C |
| ATOM | 3206 | CB | ALA | B | 196 | 59.475 | -29.787 | 3.113 | 1.00 | 31.10 | B | C |
| ATOM | 3207 | C | ALA | B | 196 | 58.422 | -30.001 | 5.369 | 1.00 | 34.20 | B | C |
| ATOM | 3208 | O | ALA | B | 196 | 59.272 | -30.672 | 5.955 | 1.00 | 34.40 | B | O |
| ATOM | 3209 | N | GLU | B | 197 | 57.643 | -29.113 | 5.980 | 1.00 | 35.33 | B | N |
| ATOM | 3210 | CA | GLU | B | 197 | 57.722 | -28.865 | 7.406 | 1.00 | 36.79 | B | C |
| ATOM | 3211 | CB | GLU | B | 197 | 58.214 | -27.436 | 7.642 | 1.00 | 39.55 | B | C |
| ATOM | 3212 | CG | GLU | B | 197 | 59.539 | -27.343 | 8.432 | 1.00 | 44.20 | B | C |
| ATOM | 3213 | CD | GLU | B | 197 | 60.699 | -28.100 | 7.791 | 1.00 | 46.09 | B | C |
| ATOM | 3214 | OE1 | GLU | B | 197 | 61.722 | -28.309 | 8.481 | 1.00 | 47.67 | B | O |
| ATOM | 3215 | OE2 | GLU | B | 197 | 60.600 | -28.481 | 6.603 | 1.00 | 48.81 | B | O |
| ATOM | 3216 | C | GLU | B | 197 | 56.339 | -29.096 | 8.023 | 1.00 | 36.65 | B | C |
| ATOM | 3217 | O | GLU | B | 197 | 56.212 | -29.783 | 9.038 | 1.00 | 36.92 | B | O |
| ATOM | 3218 | N | ASN | B | 198 | 55.301 | -28.543 | 7.399 | 1.00 | 35.53 | B | N |
| ATOM | 3219 | CA | ASN | B | 198 | 53.943 | -28.724 | 7.895 | 1.00 | 35.08 | B | C |
| ATOM | 3220 | CB | ASN | B | 198 | 53.643 | -27.717 | 8.997 | 1.00 | 36.56 | B | C |
| ATOM | 3221 | CG | ASN | B | 198 | 52.380 | -28.055 | 9.751 | 1.00 | 37.94 | B | C |
| ATOM | 3222 | OD1 | ASN | B | 198 | 51.289 | -28.084 | 9.171 | 1.00 | 38.94 | B | O |
| ATOM | 3223 | ND2 | ASN | B | 198 | 52.514 | -28.322 | 11.049 | 1.00 | 36.92 | B | N |
| ATOM | 3224 | C | ASN | B | 198 | 52.919 | -28.594 | 6.769 | 1.00 | 34.22 | B | C |
| ATOM | 3225 | O | ASN | B | 198 | 52.688 | -27.506 | 6.239 | 1.00 | 32.52 | B | O |
| ATOM | 3226 | N | ALA | B | 199 | 52.297 | -29.722 | 6.424 | 1.00 | 34.16 | B | N |
| ATOM | 3227 | CA | ALA | B | 199 | 51.326 | -29.783 | 5.340 | 1.00 | 32.28 | B | C |
| ATOM | 3228 | CB | ALA | B | 199 | 51.050 | -31.229 | 4.993 | 1.00 | 33.24 | B | C |
| ATOM | 3229 | C | ALA | B | 199 | 50.018 | -29.040 | 5.585 | 1.00 | 32.41 | B | C |
| ATOM | 3230 | O | ALA | B | 199 | 49.507 | -28.380 | 4.675 | 1.00 | 31.17 | B | O |
| ATOM | 3231 | N | GLU | B | 200 | 49.463 | -29.152 | 6.791 | 1.00 | 32.39 | B | N |

FIGURE 9-49

SUBSTITUTE SHEET (RULE 26)

59/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|--------|------|-------|---|---|
| ATOM | 3232 | CA | GLU | B | 200 | 48.217 | -28.449 | 7.106 | 1.00 | 33.82 | B | C |
| ATOM | 3233 | CB | GLU | B | 200 | 47.748 | -28.778 | 8.542 | 1.00 | 37.63 | B | C |
| ATOM | 3234 | CG | GLU | B | 200 | 46.715 | -29.924 | 8.609 | 1.00 | 43.68 | B | C |
| ATOM | 3235 | CD | GLU | B | 200 | 46.261 | -30.289 | 10.033 | 1.00 | 46.22 | B | C |
| ATOM | 3236 | OE1 | GLU | B | 200 | 47.004 | -31.015 | 10.738 | 1.00 | 47.59 | B | C |
| ATOM | 3237 | OE2 | GLU | B | 200 | 45.155 | -29.859 | 10.447 | 1.00 | 46.45 | B | O |
| ATOM | 3238 | C | GLU | B | 200 | 48.406 | -26.933 | 6.949 | 1.00 | 32.04 | B | O |
| ATOM | 3239 | O | GLU | B | 200 | 47.594 | -26.252 | 6.327 | 1.00 | 30.76 | B | O |
| ATOM | 3240 | N | VAL | B | 201 | 49.489 | -26.408 | 7.510 | 1.00 | 30.66 | B | N |
| ATOM | 3241 | CA | VAL | B | 201 | 49.764 | -24.980 | 7.422 | 1.00 | 29.45 | B | C |
| ATOM | 3242 | CB | VAL | B | 201 | 51.005 | -24.599 | 8.275 | 1.00 | 30.60 | B | C |
| ATOM | 3243 | CG1 | VAL | B | 201 | 51.337 | -23.113 | 8.088 | 1.00 | 27.46 | B | C |
| ATOM | 3244 | CG2 | VAL | B | 201 | 50.737 | -24.901 | 9.743 | 1.00 | 29.55 | B | C |
| ATOM | 3245 | C | VAL | B | 201 | 50.000 | -24.548 | 5.971 | 1.00 | 28.04 | B | C |
| ATOM | 3246 | O | VAL | B | 201 | 49.561 | -23.477 | 5.555 | 1.00 | 26.32 | B | O |
| ATOM | 3247 | N | ALA | B | 202 | 50.694 | -25.389 | 5.206 | 1.00 | 27.23 | B | N |
| ATOM | 3248 | CA | ALA | B | 202 | 50.992 | -25.085 | 3.809 | 1.00 | 26.26 | B | C |
| ATOM | 3249 | CB | ALA | B | 202 | 51.969 | -26.111 | 3.249 | 1.00 | 25.10 | B | C |
| ATOM | 3250 | C | ALA | B | 202 | 49.726 | -25.041 | 2.959 | 1.00 | 27.08 | B | C |
| ATOM | 3251 | O | ALA | B | 202 | 49.596 | -24.204 | 2.066 | 1.00 | 27.99 | B | O |
| ATOM | 3252 | N | LYS | B | 203 | 48.787 | -25.939 | 3.239 | 1.00 | 27.48 | B | N |
| ATOM | 3253 | CA | LYS | B | 203 | 47.536 | -25.969 | 2.490 | 1.00 | 28.78 | B | C |
| ATOM | 3254 | CB | LYS | B | 203 | 46.697 | -27.186 | 2.892 | 1.00 | 30.07 | B | C |
| ATOM | 3255 | CG | LYS | B | 203 | 45.575 | -27.496 | 1.916 | 1.00 | 33.81 | B | C |
| ATOM | 3256 | CD | LYS | B | 203 | 44.718 | -28.676 | 2.366 | 1.00 | 34.36 | B | C |
| ATOM | 3257 | CE | LYS | B | 203 | 43.897 | -28.334 | 3.603 | 1.00 | 36.47 | B | C |
| ATOM | 3258 | NZ | LYS | B | 203 | 43.051 | -29.489 | 4.029 | 1.00 | 38.73 | B | N |
| ATOM | 3259 | C | LYS | B | 203 | 46.728 | -24.694 | 2.744 | 1.00 | 28.51 | B | C |
| ATOM | 3260 | O | LYS | B | 203 | 46.039 | -24.198 | 1.854 | 1.00 | 27.63 | B | O |
| ATOM | 3261 | N | GLY | B | 204 | 46.830 | -24.154 | 3.954 | 1.00 | 27.86 | B | N |
| ATOM | 3262 | CA | GLY | B | 204 | 46.070 | -22.960 | 4.289 | 1.00 | 27.62 | B | C |
| ATOM | 3263 | C | GLY | B | 204 | 46.677 | -21.584 | 4.032 | 1.00 | 26.42 | B | C |
| ATOM | 3264 | O | GLY | B | 204 | 45.949 | -20.596 | 4.018 | 1.00 | 26.36 | B | O |
| ATOM | 3265 | N | VAL | B | 205 | 47.984 | -21.488 | 3.816 | 1.00 | 26.55 | B | N |
| ATOM | 3266 | CA | VAL | B | 205 | 48.589 | -20.169 | 3.601 | 1.00 | 27.10 | B | C |
| ATOM | 3267 | CB | VAL | B | 205 | 50.122 | -20.207 | 3.796 | 1.00 | 27.93 | B | C |
| ATOM | 3268 | CG1 | VAL | B | 205 | 50.747 | -21.118 | 2.763 | 1.00 | 28.76 | B | C |
| ATOM | 3269 | CG2 | VAL | B | 205 | 50.697 | -18.800 | 3.678 | 1.00 | 24.35 | B | C |
| ATOM | 3270 | C | VAL | B | 205 | 48.310 | -19.563 | 2.227 | 1.00 | 27.01 | B | C |
| ATOM | 3271 | O | VAL | B | 205 | 48.490 | -20.217 | 1.209 | 1.00 | 27.26 | B | O |
| ATOM | 3272 | N | ARG | B | 206 | 47.870 | -18.308 | 2.201 | 1.00 | 25.82 | B | N |
| ATOM | 3273 | CA | ARG | B | 206 | 47.600 | -17.638 | 0.927 | 1.00 | 26.53 | B | C |
| ATOM | 3274 | CB | ARG | B | 206 | 46.687 | -16.419 | 1.125 | 1.00 | 26.75 | B | C |
| ATOM | 3275 | CG | ARG | B | 206 | 45.214 | -16.753 | 1.374 | 1.00 | 29.27 | B | C |
| ATOM | 3276 | CD | ARG | B | 206 | 44.908 | -17.052 | 2.835 | 1.00 | 30.81 | B | C |
| ATOM | 3277 | NE | ARG | B | 206 | 44.760 | -15.832 | 3.624 | 1.00 | 30.34 | B | N |
| ATOM | 3278 | CZ | ARG | B | 206 | 43.831 | -14.901 | 3.414 | 1.00 | 31.50 | B | C |
| ATOM | 3279 | NH1 | ARG | B | 206 | 42.944 | -15.035 | 2.432 | 1.00 | 31.31 | B | N |
| ATOM | 3280 | NH2 | ARG | B | 206 | 43.794 | -13.823 | 4.184 | 1.00 | 32.17 | B | N |
| ATOM | 3281 | C | ARG | B | 206 | 48.905 | -17.185 | 0.275 | 1.00 | 24.59 | B | C |
| ATOM | 3282 | O | ARG | B | 206 | 49.770 | -16.618 | 0.936 | 1.00 | 22.91 | B | O |
| ATOM | 3283 | N | LEU | B | 207 | 49.046 | -17.442 | -1.022 | 1.00 | 24.36 | B | N |
| ATOM | 3284 | CA | LEU | B | 207 | 50.248 | -17.030 | -1.746 | 1.00 | 25.63 | B | C |
| ATOM | 3285 | CB | LEU | B | 207 | 50.970 | -18.250 | -2.326 | 1.00 | 23.92 | B | C |
| ATOM | 3286 | CG | LEU | B | 207 | 51.319 | -19.372 | -1.342 | 1.00 | 25.33 | B | C |
| ATOM | 3287 | CD1 | LEU | B | 207 | 52.076 | -20.476 | -2.076 | 1.00 | 24.00 | B | C |
| ATOM | 3288 | CD2 | LEU | B | 207 | 52.164 | -18.824 | -0.195 | 1.00 | 22.67 | B | C |
| ATOM | 3289 | C | LEU | B | 207 | 49.828 | -16.067 | -2.859 | 1.00 | 26.13 | B | C |
| ATOM | 3290 | O | LEU | B | 207 | 49.164 | -16.467 | -3.822 | 1.00 | 26.25 | B | O |
| ATOM | 3291 | N | LEU | B | 208 | 50.225 | -14.802 | -2.716 | 1.00 | 24.98 | B | N |
| ATOM | 3292 | CA | LEU | B | 208 | 49.863 | -13.753 | -3.665 | 1.00 | 24.42 | B | C |
| ATOM | 3293 | CB | LEU | B | 208 | 49.554 | -12.464 | -2.905 | 1.00 | 22.86 | B | C |
| ATOM | 3294 | CG | LEU | B | 208 | 48.660 | -12.575 | -1.665 | 1.00 | 22.52 | B | C |
| ATOM | 3295 | CD1 | LEU | B | 208 | 48.457 | -11.201 | -1.071 | 1.00 | 18.27 | B | C |
| ATOM | 3296 | CD2 | LEU | B | 208 | 47.323 | -13.207 | -2.031 | 1.00 | 22.76 | B | C |
| ATOM | 3297 | C | LEU | B | 208 | 50.903 | -13.437 | -4.741 | 1.00 | 25.81 | B | C |

FIGURE 9- 50

SUBSTITUTE SHEET (RULE 26)

60/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|---------|------|-------|---|---|
| ATOM | 3298 | O | LEU | B | 208 | 52.118 | -13.451 | -4.484 | 1.00 | 25.29 | B | O |
| ATOM | 3299 | N | TYR | B | 209 | 50.415 | -13.160 | -5.951 | 1.00 | 26.27 | B | N |
| ATOM | 3300 | CA | TYR | B | 209 | 51.295 | -12.791 | -7.046 | 1.00 | 27.62 | B | C |
| ATOM | 3301 | CB | TYR | B | 209 | 50.670 | -13.059 | -8.420 | 1.00 | 28.27 | B | C |
| ATOM | 3302 | CG | TYR | B | 209 | 51.626 | -12.712 | -9.545 | 1.00 | 29.06 | B | C |
| ATOM | 3303 | CD1 | TYR | B | 209 | 52.854 | -13.369 | -9.654 | 1.00 | 27.93 | B | C |
| ATOM | 3304 | CE1 | TYR | B | 209 | 53.783 | -13.002 | -10.618 | 1.00 | 28.72 | B | C |
| ATOM | 3305 | CD2 | TYR | B | 209 | 51.346 | -11.677 | -10.450 | 1.00 | 29.16 | B | C |
| ATOM | 3306 | CE2 | TYR | B | 209 | 52.279 | -11.302 | -11.436 | 1.00 | 28.94 | B | C |
| ATOM | 3307 | CZ | TYR | B | 209 | 53.496 | -11.971 | -11.500 | 1.00 | 30.32 | B | C |
| ATOM | 3308 | OH | TYR | B | 209 | 54.465 | -11.590 | -12.402 | 1.00 | 32.26 | B | O |
| ATOM | 3309 | C | TYR | B | 209 | 51.527 | -11.294 | -6.889 | 1.00 | 29.10 | B | C |
| ATOM | 3310 | O | TYR | B | 209 | 50.586 | -10.487 | -6.966 | 1.00 | 26.36 | B | O |
| ATOM | 3311 | N | GLY | B | 210 | 52.784 | -10.941 | -6.659 | 1.00 | 30.08 | B | N |
| ATOM | 3312 | CA | GLY | B | 210 | 53.146 | -9.552 | -6.464 | 1.00 | 32.80 | B | C |
| ATOM | 3313 | C | GLY | B | 210 | 53.784 | -8.929 | -7.677 | 1.00 | 33.52 | B | C |
| ATOM | 3314 | O | GLY | B | 210 | 54.364 | -7.855 | -7.588 | 1.00 | 34.14 | B | O |
| ATOM | 3315 | N | GLY | B | 211 | 53.688 | -9.611 | -8.812 | 1.00 | 35.13 | B | N |
| ATOM | 3316 | CA | GLY | B | 211 | 54.247 | -9.072 | -10.035 | 1.00 | 35.98 | B | C |
| ATOM | 3317 | C | GLY | B | 211 | 53.200 | -8.158 | -10.631 | 1.00 | 37.88 | B | C |
| ATOM | 3318 | O | GLY | B | 211 | 52.224 | -7.813 | -9.961 | 1.00 | 37.69 | B | O |
| ATOM | 3319 | N | SER | B | 212 | 53.373 | -7.780 | -11.889 | 1.00 | 38.64 | B | N |
| ATOM | 3320 | CA | SER | B | 212 | 52.418 | -6.890 | -12.533 | 1.00 | 41.16 | B | C |
| ATOM | 3321 | CB | SER | B | 212 | 52.854 | -6.646 | -13.982 | 1.00 | 40.83 | B | C |
| ATOM | 3322 | OG | SER | B | 212 | 52.068 | -5.635 | -14.591 | 1.00 | 41.26 | B | O |
| ATOM | 3323 | C | SER | B | 212 | 50.984 | -7.450 | -12.492 | 1.00 | 42.18 | B | C |
| ATOM | 3324 | O | SER | B | 212 | 50.730 | -8.574 | -12.938 | 1.00 | 42.76 | B | O |
| ATOM | 3325 | N | VAL | B | 213 | 50.046 | -6.672 | -11.959 | 1.00 | 42.66 | B | N |
| ATOM | 3326 | CA | VAL | B | 213 | 48.653 | -7.120 | -11.879 | 1.00 | 43.45 | B | C |
| ATOM | 3327 | CB | VAL | B | 213 | 48.237 | -7.470 | -10.420 | 1.00 | 44.23 | B | C |
| ATOM | 3328 | CG1 | VAL | B | 213 | 46.788 | -7.940 | -10.396 | 1.00 | 43.22 | B | C |
| ATOM | 3329 | CG2 | VAL | B | 213 | 49.148 | -8.542 | -9.842 | 1.00 | 44.22 | B | C |
| ATOM | 3330 | C | VAL | B | 213 | 47.687 | -6.053 | -12.382 | 1.00 | 43.88 | B | C |
| ATOM | 3331 | O | VAL | B | 213 | 47.430 | -5.069 | -11.688 | 1.00 | 43.96 | B | O |
| ATOM | 3332 | N | LYS | B | 214 | 47.159 | -6.256 | -13.587 | 1.00 | 44.72 | B | N |
| ATOM | 3333 | CA | LYS | B | 214 | 46.206 | -5.330 | -14.201 | 1.00 | 45.10 | B | C |
| ATOM | 3334 | CB | LYS | B | 214 | 46.758 | -4.793 | -15.528 | 1.00 | 45.51 | B | C |
| ATOM | 3335 | CG | LYS | B | 214 | 48.041 | -3.964 | -15.383 | 1.00 | 47.57 | B | C |
| ATOM | 3336 | CD | LYS | B | 214 | 48.433 | -3.289 | -16.697 | 1.00 | 48.72 | B | C |
| ATOM | 3337 | CE | LYS | B | 214 | 47.295 | -2.455 | -17.247 | 1.00 | 50.58 | B | C |
| ATOM | 3338 | NZ | LYS | B | 214 | 46.820 | -1.443 | -16.252 | 1.00 | 51.87 | B | N |
| ATOM | 3339 | C | LYS | B | 214 | 44.897 | -6.073 | -14.457 | 1.00 | 45.57 | B | C |
| ATOM | 3340 | O | LYS | B | 214 | 44.900 | -7.281 | -14.681 | 1.00 | 45.30 | B | O |
| ATOM | 3341 | N | ALA | B | 215 | 43.779 | -5.352 | -14.421 | 1.00 | 46.00 | B | N |
| ATOM | 3342 | CA | ALA | B | 215 | 42.475 | -5.959 | -14.646 | 1.00 | 46.04 | B | C |
| ATOM | 3343 | CB | ALA | B | 215 | 41.435 | -4.878 | -14.911 | 1.00 | 45.28 | B | C |
| ATOM | 3344 | C | ALA | B | 215 | 42.558 | -6.912 | -15.834 | 1.00 | 46.64 | B | C |
| ATOM | 3345 | O | ALA | B | 215 | 41.902 | -7.959 | -15.860 | 1.00 | 46.49 | B | O |
| ATOM | 3346 | N | ALA | B | 216 | 43.396 | -6.545 | -16.799 | 1.00 | 46.60 | B | N |
| ATOM | 3347 | CA | ALA | B | 216 | 43.588 | -7.325 | -18.020 | 1.00 | 46.13 | B | C |
| ATOM | 3348 | CB | ALA | B | 216 | 44.494 | -6.550 | -18.972 | 1.00 | 47.03 | B | C |
| ATOM | 3349 | C | ALA | B | 216 | 44.148 | -8.735 | -17.815 | 1.00 | 45.76 | B | C |
| ATOM | 3350 | O | ALA | B | 216 | 43.642 | -9.704 | -18.395 | 1.00 | 46.14 | B | O |
| ATOM | 3351 | N | SER | B | 217 | 45.190 | -8.861 | -17.004 | 1.00 | 44.20 | B | N |
| ATOM | 3352 | CA | SER | B | 217 | 45.787 | -10.175 | -16.791 | 1.00 | 44.20 | B | C |
| ATOM | 3353 | CB | SER | B | 217 | 47.286 | -10.102 | -17.084 | 1.00 | 45.35 | B | C |
| ATOM | 3354 | OG | SER | B | 217 | 47.878 | -9.010 | -16.400 | 1.00 | 46.68 | B | O |
| ATOM | 3355 | C | SER | B | 217 | 45.563 | -10.767 | -15.402 | 1.00 | 42.81 | B | C |
| ATOM | 3356 | O | SER | B | 217 | 46.309 | -11.646 | -14.961 | 1.00 | 42.42 | B | O |
| ATOM | 3357 | N | ALA | B | 218 | 44.526 | -10.298 | -14.719 | 1.00 | 41.23 | B | N |
| ATOM | 3358 | CA | ALA | B | 218 | 44.233 | -10.777 | -13.374 | 1.00 | 39.67 | B | C |
| ATOM | 3359 | CB | ALA | B | 218 | 43.355 | -9.765 | -12.648 | 1.00 | 39.24 | B | C |
| ATOM | 3360 | C | ALA | B | 218 | 43.568 | -12.153 | -13.366 | 1.00 | 38.42 | B | C |
| ATOM | 3361 | O | ALA | B | 218 | 44.071 | -13.084 | -12.733 | 1.00 | 38.19 | B | O |
| ATOM | 3362 | N | ALA | B | 219 | 42.441 | -12.271 | -14.069 | 1.00 | 36.87 | B | N |
| ATOM | 3363 | CA | ALA | B | 219 | 41.697 | -13.528 | -14.141 | 1.00 | 36.54 | B | C |

FIGURE 9 - 51

SUBSTITUTE SHEET (RULE 26)

61/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|---------|------|-------|---|---|
| ATOM | 3364 | CB | ALA | B | 219 | 40.546 | -13.405 | -15.141 | 1.00 | 34.62 | B | C |
| ATOM | 3365 | C | ALA | B | 219 | 42.592 | -14.700 | -14.517 | 1.00 | 35.47 | B | C |
| ATOM | 3366 | O | ALA | B | 219 | 42.415 | -15.799 | -14.011 | 1.00 | 35.45 | B | O |
| ATOM | 3367 | N | GLU | B | 220 | 43.552 | -14.457 | -15.402 | 1.00 | 37.32 | B | N |
| ATOM | 3368 | CA | GLU | B | 220 | 44.493 | -15.494 | -15.848 | 1.00 | 38.17 | B | C |
| ATOM | 3369 | CB | GLU | B | 220 | 45.341 | -14.977 | -17.013 | 1.00 | 41.86 | B | C |
| ATOM | 3370 | CG | GLU | B | 220 | 44.585 | -14.718 | -18.317 | 1.00 | 47.74 | B | C |
| ATOM | 3371 | CD | GLU | B | 220 | 45.464 | -14.026 | -19.357 | 1.00 | 50.73 | B | C |
| ATOM | 3372 | OE1 | GLU | B | 220 | 45.954 | -12.910 | -19.057 | 1.00 | 52.06 | B | O |
| ATOM | 3373 | OE2 | GLU | B | 220 | 45.667 | -14.589 | -20.469 | 1.00 | 52.43 | B | O |
| ATOM | 3374 | C | GLU | B | 220 | 45.439 | -15.956 | -14.733 | 1.00 | 36.23 | B | C |
| ATOM | 3375 | O | GLU | B | 220 | 45.719 | -17.153 | -14.604 | 1.00 | 35.65 | B | O |
| ATOM | 3376 | N | LEU | B | 221 | 45.954 | -15.006 | -13.952 | 1.00 | 33.56 | B | N |
| ATOM | 3377 | CA | LEU | B | 221 | 46.872 | -15.326 | -12.858 | 1.00 | 31.50 | B | C |
| ATOM | 3378 | CB | LEU | B | 221 | 47.628 | -14.068 | -12.414 | 1.00 | 32.75 | B | C |
| ATOM | 3379 | CG | LEU | B | 221 | 48.709 | -13.481 | -13.332 | 1.00 | 32.49 | B | C |
| ATOM | 3380 | CD1 | LEU | B | 221 | 48.812 | -11.965 | -13.141 | 1.00 | 32.06 | B | C |
| ATOM | 3381 | CD2 | LEU | B | 221 | 50.031 | -14.148 | -13.022 | 1.00 | 32.00 | B | C |
| ATOM | 3382 | C | LEU | B | 221 | 46.144 | -15.925 | -11.654 | 1.00 | 29.71 | B | C |
| ATOM | 3383 | O | LEU | B | 221 | 46.577 | -16.941 | -11.097 | 1.00 | 28.96 | B | O |
| ATOM | 3384 | N | PHE | B | 222 | 45.044 | -15.293 | -11.252 | 1.00 | 28.59 | B | N |
| ATOM | 3385 | CA | PHE | B | 222 | 44.279 | -15.769 | -10.102 | 1.00 | 29.46 | B | C |
| ATOM | 3386 | CB | PHE | B | 222 | 43.212 | -14.740 | -9.701 | 1.00 | 30.15 | B | C |
| ATOM | 3387 | CG | PHE | B | 222 | 43.772 | -13.383 | -9.309 | 1.00 | 31.40 | B | C |
| ATOM | 3388 | CD1 | PHE | B | 222 | 45.139 | -13.205 | -9.083 | 1.00 | 30.79 | B | C |
| ATOM | 3389 | CD2 | PHE | B | 222 | 42.924 | -12.287 | -9.150 | 1.00 | 31.09 | B | C |
| ATOM | 3390 | CE1 | PHE | B | 222 | 45.651 | -11.962 | -8.706 | 1.00 | 29.88 | B | C |
| ATOM | 3391 | CE2 | PHE | B | 222 | 43.430 | -11.043 | -8.770 | 1.00 | 29.87 | B | C |
| ATOM | 3392 | CZ | PHE | B | 222 | 44.797 | -10.884 | -8.549 | 1.00 | 30.36 | B | C |
| ATOM | 3393 | C | PHE | B | 222 | 43.624 | -17.119 | -10.392 | 1.00 | 30.54 | B | C |
| ATOM | 3394 | O | PHE | B | 222 | 43.213 | -17.840 | -9.477 | 1.00 | 29.61 | B | O |
| ATOM | 3395 | N | GLY | B | 223 | 43.542 | -17.459 | -11.674 | 1.00 | 31.13 | B | N |
| ATOM | 3396 | CA | GLY | B | 223 | 42.950 | -18.723 | -12.062 | 1.00 | 31.39 | B | C |
| ATOM | 3397 | C | GLY | B | 223 | 43.835 | -19.901 | -11.708 | 1.00 | 30.53 | B | C |
| ATOM | 3398 | O | GLY | B | 223 | 43.356 | -21.022 | -11.614 | 1.00 | 32.45 | B | O |
| ATOM | 3399 | N | MET | B | 224 | 45.126 | -19.656 | -11.501 | 1.00 | 29.21 | B | N |
| ATOM | 3400 | CA | MET | B | 224 | 46.048 | -20.731 | -11.155 | 1.00 | 26.79 | B | C |
| ATOM | 3401 | CB | MET | B | 224 | 47.496 | -20.254 | -11.294 | 1.00 | 25.00 | B | C |
| ATOM | 3402 | CG | MET | B | 224 | 47.869 | -19.867 | -12.724 | 1.00 | 26.26 | B | C |
| ATOM | 3403 | SD | MET | B | 224 | 47.701 | -21.219 | -13.943 | 1.00 | 23.16 | B | S |
| ATOM | 3404 | CE | MET | B | 224 | 49.298 | -21.693 | -14.085 | 1.00 | 20.92 | B | C |
| ATOM | 3405 | C | MET | B | 224 | 45.780 | -21.257 | -9.745 | 1.00 | 25.66 | B | C |
| ATOM | 3406 | O | MET | B | 224 | 45.493 | -20.492 | -8.826 | 1.00 | 26.03 | B | O |
| ATOM | 3407 | N | PRO | B | 225 | 45.890 | -22.582 | -9.559 | 1.00 | 24.22 | B | N |
| ATOM | 3408 | CD | PRO | B | 225 | 46.473 | -23.529 | -10.528 | 1.00 | 22.50 | B | C |
| ATOM | 3409 | CA | PRO | B | 225 | 45.652 | -23.236 | -8.266 | 1.00 | 24.99 | B | C |
| ATOM | 3410 | CB | PRO | B | 225 | 45.914 | -24.715 | -8.572 | 1.00 | 24.75 | B | C |
| ATOM | 3411 | CG | PRO | B | 225 | 46.975 | -24.650 | -9.632 | 1.00 | 25.68 | B | C |
| ATOM | 3412 | C | PRO | B | 225 | 46.454 | -22.744 | -7.051 | 1.00 | 25.52 | B | C |
| ATOM | 3413 | O | PRO | B | 225 | 45.902 | -22.629 | -5.965 | 1.00 | 25.15 | B | O |
| ATOM | 3414 | N | ASP | B | 226 | 47.741 | -22.458 | -7.223 | 1.00 | 25.22 | B | N |
| ATOM | 3415 | CA | ASP | B | 226 | 48.549 | -22.016 | -6.094 | 1.00 | 24.62 | B | C |
| ATOM | 3416 | CB | ASP | B | 226 | 49.905 | -22.725 | -6.127 | 1.00 | 24.63 | B | C |
| ATOM | 3417 | CG | ASP | B | 226 | 49.779 | -24.209 | -5.820 | 1.00 | 27.92 | B | C |
| ATOM | 3418 | OD1 | ASP | B | 226 | 49.573 | -24.560 | -4.636 | 1.00 | 28.50 | B | O |
| ATOM | 3419 | OD2 | ASP | B | 226 | 49.857 | -25.029 | -6.761 | 1.00 | 27.38 | B | O |
| ATOM | 3420 | C | ASP | B | 226 | 48.724 | -20.506 | -5.984 | 1.00 | 24.17 | B | C |
| ATOM | 3421 | O | ASP | B | 226 | 49.562 | -20.023 | -5.220 | 1.00 | 24.42 | B | O |
| ATOM | 3422 | N | ILE | B | 227 | 47.921 | -19.767 | -6.739 | 1.00 | 24.15 | B | N |
| ATOM | 3423 | CA | ILE | B | 227 | 47.957 | -18.308 | -6.703 | 1.00 | 24.90 | B | C |
| ATOM | 3424 | CB | ILE | B | 227 | 48.120 | -17.691 | -8.106 | 1.00 | 24.11 | B | C |
| ATOM | 3425 | CG2 | ILE | B | 227 | 47.932 | -16.184 | -8.038 | 1.00 | 22.64 | B | C |
| ATOM | 3426 | CG1 | ILE | B | 227 | 49.510 | -18.026 | -8.663 | 1.00 | 23.49 | B | C |
| ATOM | 3427 | CD1 | ILE | B | 227 | 49.810 | -17.387 | -10.014 | 1.00 | 23.02 | B | C |
| ATOM | 3428 | C | ILE | B | 227 | 46.630 | -17.891 | -6.101 | 1.00 | 25.85 | B | C |
| ATOM | 3429 | O | ILE | B | 227 | 45.577 | -18.014 | -6.734 | 1.00 | 26.12 | B | O |

FIGURE 9- 52

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|---------|------|-------|---|---|
| ATOM | 3430 | N | ASP | B | 228 | 46.694 | -17.402 | -4.866 | 1.00 | 26.77 | B | N |
| ATOM | 3431 | CA | ASP | B | 228 | 45.507 | -17.015 | -4.122 | 1.00 | 27.18 | B | C |
| ATOM | 3432 | CB | ASP | B | 228 | 45.700 | -17.400 | -2.654 | 1.00 | 25.65 | B | C |
| ATOM | 3433 | CG | ASP | B | 228 | 45.994 | -18.887 | -2.476 | 1.00 | 28.75 | B | C |
| ATOM | 3434 | OD1 | ASP | B | 228 | 47.011 | -19.231 | -1.825 | 1.00 | 29.36 | B | O |
| ATOM | 3435 | OD2 | ASP | B | 228 | 45.207 | -19.720 | -2.985 | 1.00 | 28.56 | B | O |
| ATOM | 3436 | C | ASP | B | 228 | 45.114 | -15.548 | -4.234 | 1.00 | 27.72 | B | C |
| ATOM | 3437 | O | ASP | B | 228 | 44.296 | -15.060 | -3.464 | 1.00 | 29.05 | B | O |
| ATOM | 3438 | N | GLY | B | 229 | 45.679 | -14.846 | -5.202 | 1.00 | 26.69 | B | N |
| ATOM | 3439 | CA | GLY | B | 229 | 45.331 | -13.449 | -5.355 | 1.00 | 26.66 | B | C |
| ATOM | 3440 | C | GLY | B | 229 | 46.560 | -12.640 | -5.682 | 1.00 | 27.08 | B | C |
| ATOM | 3441 | O | GLY | B | 229 | 47.534 | -13.164 | -6.226 | 1.00 | 27.13 | B | O |
| ATOM | 3442 | N | GLY | B | 230 | 46.536 | -11.362 | -5.339 | 1.00 | 26.39 | B | N |
| ATOM | 3443 | CA | GLY | B | 230 | 47.686 | -10.552 | -5.645 | 1.00 | 26.47 | B | C |
| ATOM | 3444 | C | GLY | B | 230 | 47.954 | -9.382 | -4.736 | 1.00 | 27.08 | B | C |
| ATOM | 3445 | O | GLY | B | 230 | 47.112 | -8.951 | -3.949 | 1.00 | 26.31 | B | O |
| ATOM | 3446 | N | LEU | B | 231 | 49.175 | -8.886 | -4.844 | 1.00 | 26.27 | B | N |
| ATOM | 3447 | CA | LEU | B | 231 | 49.595 | -7.728 | -4.088 | 1.00 | 28.33 | B | C |
| ATOM | 3448 | CB | LEU | B | 231 | 50.936 | -7.989 | -3.397 | 1.00 | 25.37 | B | C |
| ATOM | 3449 | CG | LEU | B | 231 | 51.443 | -6.777 | -2.612 | 1.00 | 24.62 | B | C |
| ATOM | 3450 | CD1 | LEU | B | 231 | 50.422 | -6.398 | -1.543 | 1.00 | 21.36 | B | C |
| ATOM | 3451 | CD2 | LEU | B | 231 | 52.791 | -7.097 | -1.995 | 1.00 | 23.54 | B | C |
| ATOM | 3452 | C | LEU | B | 231 | 49.741 | -6.689 | -5.192 | 1.00 | 27.68 | B | C |
| ATOM | 3453 | O | LEU | B | 231 | 50.755 | -6.645 | -5.881 | 1.00 | 28.70 | B | O |
| ATOM | 3454 | N | VAL | B | 232 | 48.711 | -5.874 | -5.373 | 1.00 | 28.03 | B | N |
| ATOM | 3455 | CA | VAL | B | 232 | 48.712 | -4.879 | -6.435 | 1.00 | 28.26 | B | C |
| ATOM | 3456 | CB | VAL | B | 232 | 47.262 | -4.487 | -6.813 | 1.00 | 28.04 | B | C |
| ATOM | 3457 | CG1 | VAL | B | 232 | 47.265 | -3.612 | -8.057 | 1.00 | 28.69 | B | C |
| ATOM | 3458 | CG2 | VAL | B | 232 | 46.441 | -5.731 | -7.053 | 1.00 | 29.01 | B | C |
| ATOM | 3459 | C | VAL | B | 232 | 49.506 | -3.611 | -6.135 | 1.00 | 29.08 | B | C |
| ATOM | 3460 | O | VAL | B | 232 | 49.399 | -3.028 | -5.055 | 1.00 | 30.18 | B | O |
| ATOM | 3461 | N | GLY | B | 233 | 50.296 | -3.190 | -7.114 | 1.00 | 28.95 | B | N |
| ATOM | 3462 | CA | GLY | B | 233 | 51.094 | -1.993 | -6.968 | 1.00 | 29.21 | B | C |
| ATOM | 3463 | C | GLY | B | 233 | 50.385 | -0.767 | -7.521 | 1.00 | 29.23 | B | C |
| ATOM | 3464 | O | GLY | B | 233 | 49.376 | -0.313 | -6.969 | 1.00 | 28.58 | B | O |
| ATOM | 3465 | N | GLY | B | 234 | 50.904 | -0.246 | -8.628 | 1.00 | 29.54 | B | N |
| ATOM | 3466 | CA | GLY | B | 234 | 50.335 | 0.944 | -9.241 | 1.00 | 30.30 | B | C |
| ATOM | 3467 | C | GLY | B | 234 | 48.837 | 0.934 | -9.483 | 1.00 | 31.55 | B | C |
| ATOM | 3468 | O | GLY | B | 234 | 48.154 | 1.930 | -9.230 | 1.00 | 31.14 | B | O |
| ATOM | 3469 | N | ALA | B | 235 | 48.319 | -0.189 | -9.964 | 1.00 | 30.91 | B | N |
| ATOM | 3470 | CA | ALA | B | 235 | 46.896 | -0.304 | -10.269 | 1.00 | 31.05 | B | C |
| ATOM | 3471 | CB | ALA | B | 235 | 46.587 | -1.713 | -10.768 | 1.00 | 30.62 | B | C |
| ATOM | 3472 | C | ALA | B | 235 | 45.984 | 0.051 | -9.096 | 1.00 | 31.08 | B | C |
| ATOM | 3473 | O | ALA | B | 235 | 44.871 | 0.539 | -9.298 | 1.00 | 30.40 | B | O |
| ATOM | 3474 | N | SER | B | 236 | 46.460 | -0.178 | -7.873 | 1.00 | 31.28 | B | N |
| ATOM | 3475 | CA | SER | B | 236 | 45.671 | 0.101 | -6.672 | 1.00 | 31.64 | B | C |
| ATOM | 3476 | CB | SER | B | 236 | 46.358 | -0.497 | -5.439 | 1.00 | 32.56 | B | C |
| ATOM | 3477 | OG | SER | B | 236 | 47.642 | 0.070 | -5.242 | 1.00 | 32.73 | B | O |
| ATOM | 3478 | C | SER | B | 236 | 45.397 | 1.582 | -6.418 | 1.00 | 31.55 | B | C |
| ATOM | 3479 | O | SER | B | 236 | 44.661 | 1.929 | -5.502 | 1.00 | 31.79 | B | O |
| ATOM | 3480 | N | LEU | B | 237 | 45.995 | 2.455 | -7.218 | 1.00 | 32.34 | B | N |
| ATOM | 3481 | CA | LEU | B | 237 | 45.788 | 3.888 | -7.062 | 1.00 | 33.53 | B | C |
| ATOM | 3482 | CB | LEU | B | 237 | 47.031 | 4.641 | -7.550 | 1.00 | 31.66 | B | C |
| ATOM | 3483 | CG | LEU | B | 237 | 48.294 | 4.294 | -6.748 | 1.00 | 32.78 | B | C |
| ATOM | 3484 | CD1 | LEU | B | 237 | 49.545 | 4.874 | -7.409 | 1.00 | 30.55 | B | C |
| ATOM | 3485 | CD2 | LEU | B | 237 | 48.135 | 4.820 | -5.334 | 1.00 | 29.80 | B | C |
| ATOM | 3486 | C | LEU | B | 237 | 44.547 | 4.324 | -7.854 | 1.00 | 34.47 | B | C |
| ATOM | 3487 | O | LEU | B | 237 | 44.270 | 5.517 | -7.983 | 1.00 | 34.77 | B | O |
| ATOM | 3488 | N | ASN | B | 238 | 43.809 | 3.346 | -8.377 | 1.00 | 33.75 | B | N |
| ATOM | 3489 | CA | ASN | B | 238 | 42.603 | 3.616 | -9.148 | 1.00 | 34.93 | B | C |
| ATOM | 3490 | CB | ASN | B | 238 | 42.880 | 3.403 | -10.631 | 1.00 | 35.53 | B | C |
| ATOM | 3491 | CG | ASN | B | 238 | 41.723 | 3.839 | -11.508 | 1.00 | 36.01 | B | C |
| ATOM | 3492 | OD1 | ASN | B | 238 | 40.701 | 3.149 | -11.618 | 1.00 | 34.98 | B | O |
| ATOM | 3493 | ND2 | ASN | B | 238 | 41.872 | 5.000 | -12.130 | 1.00 | 36.35 | B | N |
| ATOM | 3494 | C | ASN | B | 238 | 41.439 | 2.722 | -8.713 | 1.00 | 35.91 | B | C |
| ATOM | 3495 | O | ASN | B | 238 | 41.423 | 1.519 | -8.979 | 1.00 | 35.73 | B | O |

FIGURE 9 - 53:

SUBSTITUTE SHEET (RULE 26)

63/76

| | | | | | | | | | | | | |
|------|------|-----|-----|---|-----|--------|---------|---------|------|-------|---|---|
| ATOM | 3496 | N | ALA | B | 239 | 40.465 | 3.316 | -8.038 | 1.00 | 37.26 | B | N |
| ATOM | 3497 | CA | ALA | B | 239 | 39.300 | 2.579 | -7.563 | 1.00 | 38.65 | B | C |
| ATOM | 3498 | CB | ALA | B | 239 | 38.200 | 3.560 | -7.175 | 1.00 | 38.08 | B | C |
| ATOM | 3499 | C | ALA | B | 239 | 38.779 | 1.598 | -8.613 | 1.00 | 39.50 | B | C |
| ATOM | 3500 | O | ALA | B | 239 | 38.491 | 0.442 | -8.306 | 1.00 | 38.78 | B | O |
| ATOM | 3501 | N | ASP | B | 240 | 38.674 | 2.069 | -9.855 | 1.00 | 41.17 | B | N |
| ATOM | 3502 | CA | ASP | B | 240 | 38.166 | 1.259 | -10.963 | 1.00 | 42.15 | B | C |
| ATOM | 3503 | CB | ASP | B | 240 | 37.890 | 2.156 | -12.179 | 1.00 | 44.74 | B | C |
| ATOM | 3504 | CG | ASP | B | 240 | 36.657 | 3.035 | -11.986 | 1.00 | 48.20 | B | C |
| ATOM | 3505 | OD1 | ASP | B | 240 | 36.191 | 3.166 | -10.828 | 1.00 | 49.48 | B | O |
| ATOM | 3506 | OD2 | ASP | B | 240 | 36.156 | 3.601 | -12.989 | 1.00 | 49.33 | B | O |
| ATOM | 3507 | C | ASP | B | 240 | 39.045 | 0.079 | -11.385 | 1.00 | 41.03 | B | C |
| ATOM | 3508 | O | ASP | B | 240 | 38.544 | -1.032 | -11.535 | 1.00 | 40.18 | B | O |
| ATOM | 3509 | N | GLU | B | 241 | 40.339 | 0.297 | -11.600 | 1.00 | 39.81 | B | N |
| ATOM | 3510 | CA | GLU | B | 241 | 41.160 | -0.834 | -11.992 | 1.00 | 40.30 | B | C |
| ATOM | 3511 | CB | GLU | B | 241 | 42.555 | -0.404 | -12.448 | 1.00 | 41.71 | B | C |
| ATOM | 3512 | CG | GLU | B | 241 | 43.113 | -1.385 | -13.478 | 1.00 | 45.04 | B | C |
| ATOM | 3513 | CD | GLU | B | 241 | 44.566 | -1.128 | -13.836 | 1.00 | 48.38 | B | C |
| ATOM | 3514 | OE1 | GLU | B | 241 | 44.967 | 0.064 | -13.903 | 1.00 | 49.79 | B | O |
| ATOM | 3515 | OE2 | GLU | B | 241 | 45.302 | -2.124 | -14.069 | 1.00 | 48.00 | B | O |
| ATOM | 3516 | C | GLU | B | 241 | 41.283 | -1.817 | -10.832 | 1.00 | 38.72 | B | C |
| ATOM | 3517 | O | GLU | B | 241 | 41.313 | -3.031 | -11.041 | 1.00 | 39.89 | B | O |
| ATOM | 3518 | N | PHE | B | 242 | 41.339 | -1.292 | -9.611 | 1.00 | 36.56 | B | N |
| ATOM | 3519 | CA | PHE | B | 242 | 41.457 | -2.130 | -8.424 | 1.00 | 34.80 | B | C |
| ATOM | 3520 | CB | PHE | B | 242 | 41.592 | -1.259 | -7.172 | 1.00 | 32.39 | B | C |
| ATOM | 3521 | CG | PHE | B | 242 | 42.069 | -2.012 | -5.962 | 1.00 | 29.93 | B | C |
| ATOM | 3522 | CD1 | PHE | B | 242 | 43.355 | -2.541 | -5.923 | 1.00 | 29.06 | B | C |
| ATOM | 3523 | CD2 | PHE | B | 242 | 41.227 | -2.216 | -4.872 | 1.00 | 29.78 | B | C |
| ATOM | 3524 | CE1 | PHE | B | 242 | 43.797 | -3.262 | -4.818 | 1.00 | 26.33 | B | C |
| ATOM | 3525 | CE2 | PHE | B | 242 | 41.661 | -2.939 | -3.764 | 1.00 | 28.00 | B | C |
| ATOM | 3526 | CZ | PHE | B | 242 | 42.951 | -3.461 | -3.742 | 1.00 | 28.48 | B | C |
| ATOM | 3527 | C | PHE | B | 242 | 40.246 | -3.048 | -8.269 | 1.00 | 34.98 | B | C |
| ATOM | 3528 | O | PHE | B | 242 | 40.388 | -4.253 | -8.043 | 1.00 | 34.42 | B | O |
| ATOM | 3529 | N | GLY | B | 243 | 39.051 | -2.471 | -8.375 | 1.00 | 34.52 | B | N |
| ATOM | 3530 | CA | GLY | B | 243 | 37.843 | -3.259 | -8.244 | 1.00 | 33.64 | B | C |
| ATOM | 3531 | C | GLY | B | 243 | 37.780 | -4.362 | -9.285 | 1.00 | 33.97 | B | C |
| ATOM | 3532 | O | GLY | B | 243 | 37.322 | -5.466 | -9.004 | 1.00 | 34.21 | B | O |
| ATOM | 3533 | N | ALA | B | 244 | 38.239 | -4.065 | -10.494 | 1.00 | 33.47 | B | N |
| ATOM | 3534 | CA | ALA | B | 244 | 38.225 | -5.051 | -11.559 | 1.00 | 34.60 | B | C |
| ATOM | 3535 | CB | ALA | B | 244 | 38.714 | -4.428 | -12.850 | 1.00 | 34.04 | B | C |
| ATOM | 3536 | C | ALA | B | 244 | 39.137 | -6.199 | -11.142 | 1.00 | 35.55 | B | C |
| ATOM | 3537 | O | ALA | B | 244 | 38.825 | -7.370 | -11.359 | 1.00 | 34.72 | B | O |
| ATOM | 3538 | N | ILE | B | 245 | 40.268 | -5.869 | -10.530 | 1.00 | 35.91 | B | N |
| ATOM | 3539 | CA | ILE | B | 245 | 41.172 | -6.920 | -10.098 | 1.00 | 36.74 | B | C |
| ATOM | 3540 | CB | ILE | B | 245 | 42.512 | -6.347 | -9.585 | 1.00 | 36.73 | B | C |
| ATOM | 3541 | CG1 | ILE | B | 245 | 43.411 | -7.487 | -9.110 | 1.00 | 35.81 | B | C |
| ATOM | 3542 | CG2 | ILE | B | 245 | 43.188 | -5.549 | -10.709 | 1.00 | 36.80 | B | C |
| ATOM | 3543 | CD1 | ILE | B | 245 | 44.572 | -5.032 | -10.371 | 1.00 | 37.65 | B | C |
| ATOM | 3544 | C | ILE | B | 245 | 40.511 | -7.746 | -9.000 | 1.00 | 36.45 | B | C |
| ATOM | 3545 | O | ILE | B | 245 | 40.536 | -8.980 | -9.033 | 1.00 | 37.06 | B | O |
| ATOM | 3546 | N | CYS | B | 246 | 39.901 | -7.075 | -8.031 | 1.00 | 36.49 | B | N |
| ATOM | 3547 | CA | CYS | B | 246 | 39.249 | -7.794 | -6.945 | 1.00 | 36.70 | B | C |
| ATOM | 3548 | CB | CYS | B | 246 | 38.591 | -6.810 | -5.967 | 1.00 | 36.52 | B | C |
| ATOM | 3549 | SG | CYS | B | 246 | 39.775 | -5.752 | -5.061 | 1.00 | 36.03 | B | S |
| ATOM | 3550 | C | CYS | B | 246 | 38.208 | -8.761 | -7.502 | 1.00 | 37.42 | B | C |
| ATOM | 3551 | O | CYS | B | 246 | 38.153 | -9.929 | -7.113 | 1.00 | 37.10 | B | O |
| ATOM | 3552 | N | ARG | B | 247 | 37.383 | -8.285 | -8.425 | 1.00 | 37.90 | B | N |
| ATOM | 3553 | CA | ARG | B | 247 | 36.366 | -9.154 | -8.990 | 1.00 | 37.95 | B | C |
| ATOM | 3554 | CB | ARG | B | 247 | 35.443 | -8.372 | -9.929 | 1.00 | 39.58 | B | C |
| ATOM | 3555 | CG | ARG | B | 247 | 34.298 | -7.691 | -9.181 | 1.00 | 40.37 | B | C |
| ATOM | 3556 | CD | ARG | B | 247 | 33.245 | -7.105 | -10.120 | 1.00 | 41.08 | B | C |
| ATOM | 3557 | NE | ARG | B | 247 | 33.546 | -5.724 | -10.498 | 1.00 | 43.82 | B | N |
| ATOM | 3558 | CZ | ARG | B | 247 | 34.356 | -5.370 | -11.491 | 1.00 | 44.79 | B | C |
| ATOM | 3559 | NH1 | ARG | B | 247 | 34.954 | -6.305 | -12.222 | 1.00 | 45.37 | B | N |
| ATOM | 3560 | NH2 | ARG | B | 247 | 34.565 | -4.075 | -11.758 | 1.00 | 44.12 | B | N |
| ATOM | 3561 | C | ARG | B | 247 | 37.009 | -10.327 | -9.710 | 1.00 | 37.05 | B | C |

FIGURE 9- 54

SUBSTITUTE SHEET (RULE 26)

64/76

| | | | | | | | | | | | | |
|------|------|-----|-----|-----|-----|--------|---------|---------|------|-------|---|---|
| ATOM | 3562 | O | ARG | B | 247 | 36.510 | -11.450 | -9.639 | 1.00 | 36.84 | B | O |
| ATOM | 3563 | N | ALA | B | 248 | 38.125 | -10.080 | -10.381 | 1.00 | 35.08 | B | N |
| ATOM | 3564 | CA | ALA | B | 248 | 38.804 | -11.150 | -11.090 | 1.00 | 35.84 | B | C |
| ATOM | 3565 | CB | ALA | B | 248 | 40.084 | -10.640 | -11.727 | 1.00 | 33.58 | B | C |
| ATOM | 3566 | C | ALA | B | 248 | 39.116 | -12.292 | -10.136 | 1.00 | 35.80 | B | C |
| ATOM | 3567 | O | ALA | B | 248 | 39.158 | -13.444 | -10.542 | 1.00 | 37.62 | B | C |
| ATOM | 3568 | N | ALA | B | 249 | 39.315 | -11.978 | -8.862 | 1.00 | 35.64 | B | N |
| ATOM | 3569 | CA | ALA | B | 249 | 39.631 | -13.006 | -7.878 | 1.00 | 36.31 | B | C |
| ATOM | 3570 | CB | ALA | B | 249 | 40.381 | -12.380 | -6.704 | 1.00 | 33.52 | B | C |
| ATOM | 3571 | C | ALA | B | 249 | 38.411 | -13.769 | -7.360 | 1.00 | 37.06 | B | C |
| ATOM | 3572 | O | ALA | B | 249 | 38.549 | -14.793 | -6.702 | 1.00 | 36.02 | B | O |
| ATOM | 3573 | N | GLY | B | 250 | 37.217 | -13.273 | -7.650 | 1.00 | 39.46 | B | N |
| ATOM | 3574 | CA | GLY | B | 250 | 36.026 | -13.944 | -7.164 | 1.00 | 42.33 | B | C |
| ATOM | 3575 | C | GLY | B | 250 | 35.567 | -15.111 | -8.014 | 1.00 | 44.29 | B | C |
| ATOM | 3576 | O | GLY | B | 250 | 34.335 | -15.230 | -8.226 | 1.00 | 45.84 | B | O |
| ATOM | 3577 | OXT | GLY | B | 250 | 36.427 | -15.917 | -8.448 | 1.00 | 44.68 | B | O |
| TER | 3578 | GLY | B | 250 | | | | | | | B | O |
| ATOM | 3579 | OH2 | WAT | | 1 | 63.739 | 13.680 | 3.442 | 1.00 | 21.18 | S | |
| ATOM | 3580 | OH2 | WAT | | 2 | 46.539 | 11.585 | 11.681 | 1.00 | 18.96 | S | |
| ATOM | 3581 | OH2 | WAT | | 3 | 43.567 | -19.385 | -7.432 | 1.00 | 25.91 | S | |
| ATOM | 3582 | OH2 | WAT | | 4 | 71.012 | 30.393 | 25.191 | 1.00 | 32.51 | S | |
| ATOM | 3583 | OH2 | WAT | | 5 | 75.981 | 26.094 | 0.757 | 1.00 | 20.73 | S | |
| ATOM | 3584 | OH2 | WAT | | 6 | 48.777 | 11.393 | 9.884 | 1.00 | 22.78 | S | |
| ATOM | 3585 | OH2 | WAT | | 7 | 56.505 | -5.204 | 12.895 | 1.00 | 19.66 | S | |
| ATOM | 3586 | OH2 | WAT | | 8 | 59.027 | 10.900 | 23.449 | 1.00 | 22.67 | S | |
| ATOM | 3587 | OH2 | WAT | | 9 | 53.790 | -3.883 | -4.465 | 1.00 | 19.70 | S | |
| ATOM | 3588 | OH2 | WAT | | 10 | 55.768 | -13.300 | 17.134 | 1.00 | 24.11 | S | |
| ATOM | 3589 | OH2 | WAT | | 11 | 55.028 | 18.137 | 22.898 | 1.00 | 25.58 | S | |
| ATOM | 3590 | OH2 | WAT | | 12 | 72.703 | 11.819 | 16.372 | 1.00 | 22.99 | S | |
| ATOM | 3591 | OH2 | WAT | | 13 | 49.271 | 16.075 | 28.292 | 1.00 | 27.82 | S | |
| ATOM | 3592 | OH2 | WAT | | 14 | 45.485 | 10.010 | 8.553 | 1.00 | 25.65 | S | |
| ATOM | 3593 | OH2 | WAT | | 15 | 59.531 | -0.501 | 22.088 | 1.00 | 32.02 | S | |
| ATOM | 3594 | OH2 | WAT | | 16 | 56.843 | 11.447 | 12.086 | 1.00 | 26.98 | S | |
| ATOM | 3595 | OH2 | WAT | | 17 | 61.865 | 0.446 | 9.105 | 1.00 | 25.39 | S | |
| ATOM | 3596 | OH2 | WAT | | 18 | 58.552 | -3.361 | 22.182 | 1.00 | 21.78 | S | |
| ATOM | 3597 | OH2 | WAT | | 19 | 82.120 | 18.194 | 26.672 | 1.00 | 27.91 | S | |
| ATOM | 3598 | OH2 | WAT | | 20 | 46.996 | -16.906 | 5.047 | 1.00 | 31.56 | S | |
| ATOM | 3599 | OH2 | WAT | | 21 | 57.610 | -4.027 | 3.628 | 1.00 | 23.68 | S | |
| ATOM | 3600 | OH2 | WAT | | 22 | 52.890 | -0.095 | 23.500 | 1.00 | 32.40 | S | |
| ATOM | 3601 | OH2 | WAT | | 23 | 47.078 | 8.133 | 9.748 | 1.00 | 25.02 | S | |
| ATOM | 3602 | OH2 | WAT | | 24 | 65.725 | 10.982 | 12.443 | 1.00 | 21.03 | S | |
| ATOM | 3603 | OH2 | WAT | | 25 | 68.888 | 27.520 | 12.697 | 1.00 | 30.16 | S | |
| ATOM | 3604 | OH2 | WAT | | 26 | 72.959 | 24.629 | -0.833 | 1.00 | 28.76 | S | |
| ATOM | 3605 | OH2 | WAT | | 27 | 64.407 | 24.946 | 34.919 | 1.00 | 30.03 | S | |
| ATOM | 3606 | OH2 | WAT | | 28 | 82.109 | 28.612 | 19.089 | 1.00 | 25.80 | S | |
| ATOM | 3607 | OH2 | WAT | | 29 | 59.333 | 24.901 | 6.708 | 1.00 | 29.91 | S | |
| ATOM | 3608 | OH2 | WAT | | 30 | 60.572 | -8.956 | -1.830 | 1.00 | 24.69 | S | |
| ATOM | 3609 | OH2 | WAT | | 31 | 71.583 | 13.346 | 28.698 | 1.00 | 31.02 | S | |
| ATOM | 3610 | OH2 | WAT | | 32 | 81.021 | 32.930 | 13.694 | 1.00 | 37.30 | S | |
| ATOM | 3611 | OH2 | WAT | | 33 | 50.899 | 12.898 | 29.572 | 1.00 | 30.82 | S | |
| ATOM | 3612 | OH2 | WAT | | 34 | 42.795 | 14.907 | 18.385 | 1.00 | 28.46 | S | |
| ATOM | 3613 | OH2 | WAT | | 35 | 75.426 | 12.778 | -9.279 | 1.00 | 29.79 | S | |
| ATOM | 3614 | OH2 | WAT | | 36 | 53.891 | 23.103 | 10.722 | 1.00 | 26.52 | S | |
| ATOM | 3615 | OH2 | WAT | | 37 | 43.505 | 0.281 | -3.618 | 1.00 | 32.25 | S | |
| ATOM | 3616 | OH2 | WAT | | 38 | 51.056 | 11.344 | 25.909 | 1.00 | 34.32 | S | |
| ATOM | 3617 | OH2 | WAT | | 39 | 63.457 | 5.623 | 6.569 | 1.00 | 26.78 | S | |
| ATOM | 3618 | OH2 | WAT | | 40 | 42.887 | 5.137 | 21.307 | 1.00 | 36.91 | S | |
| ATOM | 3619 | OH2 | WAT | | 41 | 63.977 | 6.391 | 2.587 | 1.00 | 31.27 | S | |
| ATOM | 3620 | OH2 | WAT | | 42 | 62.463 | -0.901 | 12.868 | 1.00 | 26.15 | S | |
| ATOM | 3621 | OH2 | WAT | | 43 | 63.052 | -9.913 | -2.660 | 1.00 | 31.73 | S | |
| ATOM | 3622 | OH2 | WAT | | 44 | 84.960 | 22.849 | 9.750 | 1.00 | 32.28 | S | |
| ATOM | 3623 | OH2 | WAT | | 45 | 53.936 | 6.657 | 12.882 | 1.00 | 36.99 | S | |
| ATOM | 3624 | OH2 | WAT | | 46 | 66.545 | 30.127 | 26.084 | 1.00 | 34.75 | S | |
| ATOM | 3625 | OH2 | WAT | | 47 | 83.758 | 32.772 | 7.138 | 1.00 | 31.70 | S | |
| ATOM | 3626 | OH2 | WAT | | 48 | 55.487 | 14.405 | -4.745 | 1.00 | 27.79 | S | |
| ATOM | 3627 | OH2 | WAT | | 49 | 68.567 | 19.860 | -1.275 | 1.00 | 25.45 | S | |

FIGURE 9 - 55

SUBSTITUTE SHEET (RULE 26)

65/76

| | | | | | | | | | | |
|------|------|-----|-----|-----|--------|---------|---------|------|-------|---|
| ATOM | 3628 | OH2 | WAT | 50 | 35.299 | -8.627 | -1.566 | 1.00 | 37.22 | S |
| ATOM | 3629 | OH2 | WAT | 51 | 75.282 | 7.386 | 10.848 | 1.00 | 28.73 | S |
| ATOM | 3630 | OH2 | WAT | 52 | 47.602 | 21.845 | 29.388 | 1.00 | 36.21 | S |
| ATOM | 3631 | OH2 | WAT | 53 | 56.100 | 29.843 | 22.330 | 1.00 | 29.84 | S |
| ATOM | 3632 | OH2 | WAT | 54 | 70.307 | 21.808 | 33.869 | 1.00 | 35.11 | S |
| ATOM | 3633 | OH2 | WAT | 55 | 51.842 | 2.180 | 5.586 | 1.00 | 27.16 | S |
| ATOM | 3634 | OH2 | WAT | 56 | 49.428 | 18.363 | 30.767 | 1.00 | 34.50 | S |
| ATOM | 3635 | OH2 | WAT | 57 | 52.432 | -26.252 | -10.500 | 1.00 | 31.73 | S |
| ATOM | 3636 | OH2 | WAT | 58 | 88.006 | 14.325 | 18.106 | 1.00 | 38.40 | S |
| ATOM | 3637 | OH2 | WAT | 59 | 56.307 | 28.467 | 17.331 | 1.00 | 35.26 | S |
| ATOM | 3638 | OH2 | WAT | 60 | 72.612 | 30.570 | 9.885 | 1.00 | 30.94 | S |
| ATOM | 3639 | OH2 | WAT | 61 | 59.289 | -25.133 | 11.886 | 1.00 | 36.57 | S |
| ATOM | 3640 | OH2 | WAT | 62 | 62.849 | -12.930 | 4.372 | 1.00 | 34.28 | S |
| ATOM | 3641 | OH2 | WAT | 63 | 48.208 | 11.640 | 26.364 | 1.00 | 31.54 | S |
| ATOM | 3642 | OH2 | WAT | 64 | 57.116 | -25.895 | -8.962 | 1.00 | 28.63 | S |
| ATOM | 3643 | OH2 | WAT | 65 | 57.543 | 3.069 | 8.138 | 1.00 | 36.57 | S |
| ATOM | 3644 | OH2 | WAT | 66 | 51.001 | 2.103 | -5.920 | 1.00 | 27.74 | S |
| ATOM | 3645 | OH2 | WAT | 67 | 61.087 | -8.616 | 16.804 | 1.00 | 33.04 | S |
| ATOM | 3646 | OH2 | WAT | 68 | 66.425 | -6.837 | -11.132 | 1.00 | 39.35 | S |
| ATOM | 3647 | OH2 | WAT | 69 | 77.598 | 22.382 | 34.696 | 1.00 | 34.40 | S |
| ATOM | 3648 | OH2 | WAT | 70 | 54.065 | 4.324 | 5.269 | 1.00 | 32.40 | S |
| ATOM | 3649 | OH2 | WAT | 71 | 48.065 | -21.741 | 7.107 | 1.00 | 28.43 | S |
| ATOM | 3650 | OH2 | WAT | 72 | 77.492 | 32.093 | 20.572 | 1.00 | 32.41 | S |
| ATOM | 3651 | OH2 | WAT | 73 | 67.956 | 17.031 | 32.792 | 1.00 | 38.47 | S |
| ATOM | 3652 | OH2 | WAT | 74 | 35.792 | -1.645 | -11.357 | 1.00 | 41.03 | S |
| ATOM | 3653 | OH2 | WAT | 75 | 72.097 | 31.059 | 22.915 | 1.00 | 31.39 | S |
| ATOM | 3654 | OH2 | WAT | 76 | 60.828 | -5.551 | 5.997 | 1.00 | 35.88 | S |
| ATOM | 3655 | OH2 | WAT | 77 | 86.099 | 14.188 | 11.173 | 1.00 | 36.28 | S |
| ATOM | 3656 | OH2 | WAT | 78 | 74.845 | 10.187 | 16.911 | 1.00 | 30.97 | S |
| ATOM | 3657 | OH2 | WAT | 79 | 50.465 | 32.604 | 25.734 | 1.00 | 35.32 | S |
| ATOM | 3658 | OH2 | WAT | 80 | 42.236 | -11.831 | 5.496 | 1.00 | 38.33 | S |
| ATOM | 3659 | OH2 | WAT | 81 | 58.627 | -4.082 | 6.362 | 1.00 | 34.82 | S |
| ATOM | 3660 | OH2 | WAT | 82 | 67.975 | -11.445 | -2.979 | 1.00 | 36.42 | S |
| ATOM | 3661 | OH2 | WAT | 83 | 45.950 | 8.709 | -4.339 | 1.00 | 41.88 | S |
| ATOM | 3662 | OH2 | WAT | 84 | 66.646 | 16.717 | -3.564 | 1.00 | 38.92 | S |
| ATOM | 3663 | OH2 | WAT | 85 | 76.878 | 19.131 | -4.499 | 1.00 | 30.97 | S |
| ATOM | 3664 | OH2 | WAT | 86 | 68.904 | 30.992 | 21.253 | 1.00 | 28.40 | S |
| ATOM | 3665 | OH2 | WAT | 87 | 78.208 | 8.520 | 16.860 | 1.00 | 28.01 | S |
| ATOM | 3666 | OH2 | WAT | 88 | 54.448 | 31.296 | 29.005 | 1.00 | 44.80 | S |
| ATOM | 3667 | OH2 | WAT | 89 | 67.347 | 24.984 | 35.233 | 1.00 | 34.81 | S |
| ATOM | 3668 | OH2 | WAT | 90 | 60.370 | -26.116 | -2.606 | 1.00 | 35.93 | S |
| ATOM | 3669 | OH2 | WAT | 91 | 37.567 | 1.014 | 8.342 | 1.00 | 33.11 | S |
| ATOM | 3670 | OH2 | WAT | 92 | 65.188 | 12.783 | 21.098 | 1.00 | 26.41 | S |
| ATOM | 3671 | OH2 | WAT | 93 | 65.456 | 11.831 | 23.641 | 1.00 | 34.38 | S |
| ATOM | 3672 | OH2 | WAT | 94 | 56.795 | -30.673 | 0.572 | 1.00 | 42.38 | S |
| ATOM | 3673 | OH2 | WAT | 95 | 56.901 | 8.116 | 12.364 | 1.00 | 38.07 | S |
| ATOM | 3674 | OH2 | WAT | 96 | 42.831 | 16.110 | 15.976 | 1.00 | 34.52 | S |
| ATOM | 3675 | OH2 | WAT | 97 | 42.170 | 16.768 | 20.681 | 1.00 | 28.67 | S |
| ATOM | 3676 | OH2 | WAT | 98 | 57.637 | 7.463 | 14.873 | 1.00 | 31.48 | S |
| ATOM | 3677 | OH2 | WAT | 99 | 56.202 | 35.047 | 27.703 | 1.00 | 32.52 | S |
| ATOM | 3678 | OH2 | WAT | 100 | 41.686 | -3.364 | 9.116 | 1.00 | 28.74 | S |
| ATOM | 3679 | OH2 | WAT | 101 | 86.586 | 3.817 | 23.769 | 1.00 | 41.10 | S |
| ATOM | 3680 | OH2 | WAT | 102 | 58.302 | 1.697 | 11.346 | 1.00 | 31.83 | S |
| ATOM | 3681 | OH2 | WAT | 103 | 58.832 | 34.722 | 27.087 | 1.00 | 38.83 | S |
| ATOM | 3682 | OH2 | WAT | 104 | 85.938 | 9.832 | 11.193 | 1.00 | 46.57 | S |
| ATOM | 3683 | OH2 | WAT | 105 | 41.407 | 11.430 | 24.565 | 1.00 | 32.31 | S |
| ATOM | 3684 | OH2 | WAT | 106 | 60.136 | 7.850 | 5.330 | 1.00 | 30.07 | S |
| ATOM | 3685 | OH2 | WAT | 107 | 86.102 | 28.442 | 14.054 | 1.00 | 34.85 | S |
| ATOM | 3686 | OH2 | WAT | 108 | 53.305 | 24.445 | 31.958 | 1.00 | 39.87 | S |
| ATOM | 3687 | OH2 | WAT | 109 | 65.921 | 10.467 | 9.630 | 1.00 | 34.53 | S |
| ATOM | 3688 | OH2 | WAT | 110 | 68.629 | 4.220 | 1.225 | 1.00 | 45.83 | S |
| ATOM | 3689 | OH2 | WAT | 111 | 68.562 | 10.278 | 8.956 | 1.00 | 46.62 | S |
| ATOM | 3690 | OH2 | WAT | 112 | 56.017 | -15.270 | 15.035 | 1.00 | 36.72 | S |
| ATOM | 3691 | OH2 | WAT | 113 | 59.462 | 26.860 | 33.469 | 1.00 | 35.19 | S |
| ATOM | 3692 | OH2 | WAT | 114 | 60.433 | 9.337 | -5.622 | 1.00 | 26.88 | S |
| ATOM | 3693 | OH2 | WAT | 115 | 66.775 | -10.262 | 0.003 | 1.00 | 42.01 | S |

FIGURE 9- 56

SUBSTITUTE SHEET (RULE 26)

| | | | | | | | | | | |
|------|------|-----|-----|-----|--------|---------|---------|------|-------|---|
| ATOM | 3694 | OH2 | WAT | 116 | 67.236 | 14.471 | 21.708 | 1.00 | 39.77 | S |
| ATOM | 3695 | OH2 | WAT | 117 | 79.784 | 3.290 | 23.426 | 1.00 | 38.65 | S |
| ATOM | 3696 | OH2 | WAT | 118 | 57.689 | 2.771 | 26.058 | 1.00 | 35.29 | S |
| ATOM | 3697 | OH2 | WAT | 119 | 55.846 | -24.062 | 14.597 | 1.00 | 40.39 | S |
| ATOM | 3698 | OH2 | WAT | 120 | 70.641 | 6.345 | 21.248 | 1.00 | 36.46 | S |
| ATOM | 3699 | OH2 | WAT | 121 | 78.813 | 15.534 | 34.365 | 1.00 | 30.13 | S |
| ATOM | 3700 | OH2 | WAT | 122 | 71.200 | 16.610 | 30.358 | 1.00 | 42.54 | S |
| ATOM | 3701 | OH2 | WAT | 123 | 70.119 | -11.382 | 7.901 | 1.00 | 35.98 | S |
| ATOM | 3702 | OH2 | WAT | 124 | 69.454 | 7.610 | 29.487 | 1.00 | 45.35 | S |
| ATOM | 3703 | OH2 | WAT | 125 | 66.807 | 32.983 | 14.967 | 1.00 | 32.20 | S |
| ATOM | 3704 | OH2 | WAT | 126 | 59.180 | 6.165 | -8.206 | 1.00 | 32.59 | S |
| ATOM | 3705 | OH2 | WAT | 127 | 67.208 | -1.417 | 10.324 | 1.00 | 35.14 | S |
| ATOM | 3706 | OH2 | WAT | 128 | 53.773 | -6.675 | -5.077 | 1.00 | 40.92 | S |
| ATOM | 3707 | OH2 | WAT | 129 | 37.999 | 9.731 | 6.178 | 1.00 | 37.69 | S |
| ATOM | 3708 | OH2 | WAT | 130 | 70.760 | 31.115 | -1.126 | 1.00 | 48.87 | S |
| ATOM | 3709 | OH2 | WAT | 131 | 48.587 | 15.254 | 32.305 | 1.00 | 42.90 | S |
| ATOM | 3710 | OH2 | WAT | 132 | 46.698 | 17.046 | 28.377 | 1.00 | 33.90 | S |
| ATOM | 3711 | OH2 | WAT | 133 | 69.756 | -10.705 | 4.417 | 1.00 | 40.92 | S |
| ATOM | 3712 | OH2 | WAT | 134 | 61.801 | 11.765 | -5.827 | 1.00 | 30.08 | S |
| ATOM | 3713 | OH2 | WAT | 135 | 65.150 | 33.908 | 25.255 | 1.00 | 40.06 | S |
| ATOM | 3714 | OH2 | WAT | 136 | 64.483 | 23.992 | 38.822 | 1.00 | 37.47 | S |
| ATOM | 3715 | OH2 | WAT | 137 | 68.996 | -15.404 | 5.846 | 1.00 | 43.32 | S |
| ATOM | 3716 | OH2 | WAT | 138 | 81.024 | 20.065 | 34.384 | 1.00 | 35.58 | S |
| ATOM | 3717 | OH2 | WAT | 139 | 87.308 | 15.441 | 20.501 | 1.00 | 45.36 | S |
| ATOM | 3718 | OH2 | WAT | 140 | 77.132 | 5.263 | 24.102 | 1.00 | 38.49 | S |
| ATOM | 3719 | OH2 | WAT | 141 | 36.486 | -16.779 | -11.072 | 1.00 | 38.24 | S |
| ATOM | 3720 | OH2 | WAT | 142 | 84.757 | 23.777 | 23.127 | 1.00 | 37.17 | S |
| ATOM | 3721 | OH2 | WAT | 143 | 70.264 | 28.931 | 10.768 | 1.00 | 31.00 | S |
| ATOM | 3722 | OH2 | WAT | 144 | 41.269 | -10.622 | -16.396 | 1.00 | 41.13 | S |
| ATOM | 3723 | OH2 | WAT | 145 | 68.274 | 7.143 | 22.602 | 1.00 | 46.06 | S |
| ATOM | 3724 | OH2 | WAT | 146 | 62.567 | 14.392 | 28.562 | 1.00 | 44.69 | S |
| ATOM | 3725 | OH2 | WAT | 147 | 64.731 | 24.523 | 4.703 | 1.00 | 36.32 | S |
| ATOM | 3726 | OH2 | WAT | 148 | 43.669 | -20.914 | -5.240 | 1.00 | 38.93 | S |
| ATOM | 3727 | OH2 | WAT | 149 | 34.494 | 14.494 | -5.873 | 1.00 | 40.21 | S |
| ATOM | 3728 | OH2 | WAT | 150 | 63.549 | -12.958 | 6.999 | 1.00 | 32.43 | S |
| ATOM | 3729 | OH2 | WAT | 151 | 59.144 | 32.488 | 29.232 | 1.00 | 44.36 | S |
| ATOM | 3730 | OH2 | WAT | 152 | 64.594 | 12.517 | 7.745 | 1.00 | 43.80 | S |
| ATOM | 3731 | OH2 | WAT | 153 | 42.616 | -18.758 | -3.095 | 1.00 | 36.69 | S |
| ATOM | 3732 | OH2 | WAT | 154 | 54.449 | 4.642 | 2.609 | 1.00 | 38.54 | S |
| ATOM | 3733 | OH2 | WAT | 155 | 50.417 | 0.550 | 21.384 | 1.00 | 44.81 | S |
| ATOM | 3734 | OH2 | WAT | 156 | 37.279 | -8.043 | -13.535 | 1.00 | 38.19 | S |
| ATOM | 3735 | OH2 | WAT | 157 | 49.298 | -13.025 | 12.873 | 1.00 | 37.47 | S |
| ATOM | 3736 | OH2 | WAT | 158 | 41.030 | 24.450 | 22.675 | 1.00 | 46.89 | S |
| ATOM | 3737 | OH2 | WAT | 159 | 72.495 | 2.930 | 6.308 | 1.00 | 38.50 | S |
| ATOM | 3738 | OH2 | WAT | 160 | 61.298 | 27.611 | 13.565 | 1.00 | 46.96 | S |
| ATOM | 3739 | OH2 | WAT | 161 | 62.971 | 36.604 | 18.041 | 1.00 | 44.82 | S |
| ATOM | 3740 | OH2 | WAT | 162 | 83.572 | 13.349 | 29.985 | 1.00 | 43.32 | S |
| ATOM | 3741 | OH2 | WAT | 163 | 36.643 | 10.222 | 1.115 | 1.00 | 29.66 | S |
| ATOM | 3742 | OH2 | WAT | 164 | 57.618 | -4.939 | -7.098 | 1.00 | 41.52 | S |
| ATOM | 3743 | OH2 | WAT | 165 | 77.841 | 8.652 | 12.101 | 1.00 | 34.90 | S |
| ATOM | 3744 | OH2 | WAT | 166 | 40.961 | -15.929 | -6.954 | 1.00 | 44.03 | S |
| ATOM | 3745 | OH2 | WAT | 167 | 66.564 | -9.125 | 2.563 | 1.00 | 36.50 | S |
| ATOM | 3746 | OH2 | WAT | 168 | 57.004 | -4.662 | 8.557 | 1.00 | 30.72 | S |
| ATOM | 3747 | OH2 | WAT | 169 | 43.288 | -20.287 | 4.714 | 1.00 | 47.63 | S |
| ATOM | 3748 | OH2 | WAT | 170 | 75.464 | 26.284 | 35.459 | 1.00 | 44.63 | S |
| ATOM | 3749 | OH2 | WAT | 171 | 53.679 | 5.266 | 25.380 | 1.00 | 45.04 | S |
| ATOM | 3750 | OH2 | WAT | 172 | 56.443 | 12.148 | -7.782 | 1.00 | 31.88 | S |
| ATOM | 3751 | OH2 | WAT | 173 | 70.455 | -16.383 | -1.712 | 1.00 | 44.98 | S |
| ATOM | 3752 | OH2 | WAT | 174 | 30.948 | -3.393 | 2.233 | 1.00 | 40.95 | S |
| ATOM | 3753 | OH2 | WAT | 175 | 41.711 | 20.576 | 18.327 | 1.00 | 43.75 | S |
| ATOM | 3754 | OH2 | WAT | 176 | 78.523 | 5.879 | 7.675 | 1.00 | 42.69 | S |
| ATOM | 3755 | OH2 | WAT | 177 | 38.101 | 5.062 | -10.304 | 1.00 | 40.30 | S |
| ATOM | 3756 | OH2 | WAT | 178 | 54.268 | 29.952 | 20.197 | 1.00 | 35.72 | S |
| ATOM | 3757 | OH2 | WAT | 179 | 72.440 | 29.012 | 32.447 | 1.00 | 39.75 | S |
| ATOM | 3758 | OH2 | WAT | 180 | 80.516 | 13.435 | 32.834 | 1.00 | 43.98 | S |
| ATOM | 3759 | OH2 | WAT | 181 | 64.768 | 1.119 | 21.506 | 1.00 | 43.72 | S |

FIGURE 9- 57

SUBSTITUTE SHEET (RULE 26)

67/76

| | | | | | | | | | |
|------|------|---------|-----|--------|---------|--------|------|-------|---|
| ATOM | 3760 | OH2 WAT | 182 | 86.668 | 13.793 | 22.641 | 1.00 | 45.55 | S |
| ATOM | 3761 | OH2 WAT | 183 | 53.783 | 15.211 | -2.617 | 1.00 | 41.77 | S |
| ATOM | 3762 | OH2 WAT | 184 | 47.017 | 17.986 | 7.210 | 1.00 | 47.29 | S |
| ATOM | 3763 | OH2 WAT | 185 | 47.776 | -19.055 | 6.655 | 1.00 | 47.45 | S |
| ATOM | 3764 | OH2 WAT | 186 | 61.369 | 35.933 | 24.964 | 1.00 | 43.41 | S |
| ATOM | 3765 | OH2 WAT | 187 | 55.247 | -3.515 | -6.768 | 1.00 | 40.64 | S |
| ATOM | 3766 | OH2 WAT | 188 | 42.534 | 9.618 | -6.369 | 1.00 | 39.21 | S |
| ATOM | 3767 | OH2 WAT | 189 | 55.186 | 5.737 | 16.154 | 1.00 | 41.22 | S |
| ATOM | 3768 | OH2 WAT | 190 | 68.693 | 23.313 | 36.902 | 1.00 | 39.71 | S |
| ATOM | 3769 | OH2 WAT | 191 | 78.189 | 7.398 | 4.865 | 1.00 | 47.98 | S |
| ATOM | 3770 | OH2 WAT | 192 | 45.493 | -14.622 | 6.823 | 1.00 | 37.56 | S |
| ATOM | 3771 | OH2 WAT | 193 | 77.993 | 11.404 | -1.840 | 1.00 | 41.99 | S |
| ATOM | 3772 | OH2 WAT | 194 | 37.218 | -19.477 | 0.511 | 1.00 | 47.19 | S |

FIGURE 9-58

SUBSTITUTE SHEET (RULE 26)

FIGURE 10

```

EC_tpiA 1 -----
PA_tpiA 1 -----
SP_tpiA 1 -----
EF_tpiA 1 -----
SA_tpiA 1 -----
HP_tpiA 1 -----

EC_tpiA 28 KEIAGVAGCAVSIKPPFEDIMAKREAE-----SHMLGAQVVDLN-LSGAFTGETSAAAI
PA_tpiA 28 -QALPSGVIVVMPFCLEISQIQGLAE-----KARDGAQNSAVSPMOSATIGETTESOI
SP_tpiA 36 SKLESSDLTPAGISAPALDITTLAVAKG-----SNKVAQMCYNE-NAGAFGETSPQVLE
EF_tpiA 28 NAFESNDVVPVAVISSPALMLAPAWNED-----SEKLEAQCYNH-NAGAFGETGENSPAAI
SA_tpiA 61 -TLEDSEKVESVFCAPFQIDATTAKEGKAQGLGGAQRTYHE-DNGAFGETGETSEVAL
HP_tpiA 27 KTKPKQHPDRNTPFEDFFGLPNSFLHFT-----LGVDHAYPR-DCSAFTGETTSKHL

EC_tpiA 84 KDHCAQYVHIGHSERRYTHDEDELIAKKFAVERKEGHTPLACGGETPAENSACKTEEVC
PA_tpiA 84 ADNGCSMVHIGHSERRLLILGDEVYSKKFAAQSCSVVPSICGSETRAEAGKTLZVV
SP_tpiA 92 KETGTDYVHIGHSERRYFHEDEHINKKKAIFAMGMLPAHCCSEKLETYEAGKTADEV
EF_tpiA 84 ADLGNDYVHIGHSERRYFHEDEHINKKKAIFAMGMLPAHCCSEKLETYEAGKTADEV
SA_tpiA 119 ADLGKDYVHIGHSERRLLFHEDEHINKKKAIFAMGMLPAHCCSEKLETYEAGKTADEV
HP_tpiA 79 EELKHTLHIGHSERRLLFHEDEHINKKKAIFAMGMLPAHCCSEKLETYEAGKTADEV

EC_tpiA 144 PRGDAVKTQGAFAFEGAVIAYEPIWAIGTGKSSAPACQAPHKFEDHAKKD-ANEA
PA_tpiA 144 PRGGSVIDEYGVGAFERAVIAYEPIWAIGTGLASPAQAOEHHFAMBAQNAEN-NEVA
SP_tpiA 152 PAKSANTAGLPAEQVRSVIAYEPIWAIGTGKSASQEDACKCKVVEDVVAADFGQEVA
EF_tpiA 144 EGCHKKMLVGLNNEGVSESVIAYEPIWAIGTGKSADANIEDECCSVVSTREKLYGKEVS
SA_tpiA 179 ECKKKVAGLSEEQKLSVIAYEPIWAIGTGKSMSIDANEHCHEVROTADSSKEVS
HP_tpiA 137 KEFLSEOLENIDLN-YPNLIWVAYEPIWAIGTGKSASLEDEIYLTHGFEKQIENQRT-----

EC_tpiA 203 ECVIIQYGGSVNASNAELFQPDLDGALVGGASIKALPAEAVLYKAEAAKQA
PA_tpiA 203 KVVRIIYGGSVNAASNAELFQPDLDGALVGGASINAEPCACCRAGS-----
SP_tpiA 212 EKVRIIYGGSVKPEHVAASYACPDLDGALVGGASLEADSPAPLIFVY-----
EF_tpiA 204 EKVRIQYGGSVKPEHIAEYNKKEPDLDGALVGGASLEADSPAPLIFVY-----
SA_tpiA 239 EKVRIQYGGSVKPNNIKEYMAOTPDLDGALVGGASLKVDETHOLEBAK-----
HP_tpiA 191 ---ELKGGSVNTQNAKEITLIDSADGILGGSASLEADSPAPLIFVY-----

```

Organisms are abbreviated as follows: EC = *Escherichia coli*; HP = *Helicobacter pylori*; PA = *Pseudomonas aeruginosa*; SA = *Staphylococcus aureus*; SP = *Streptococcus pneumoniae*; EF = *Enterococcus faecalis*.

FIGURE 11

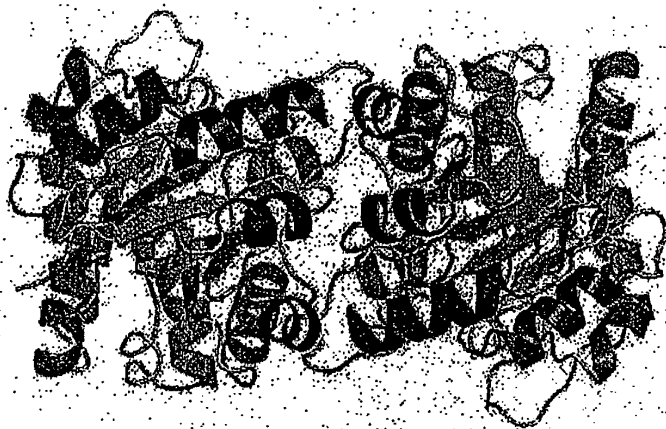


FIGURE 12

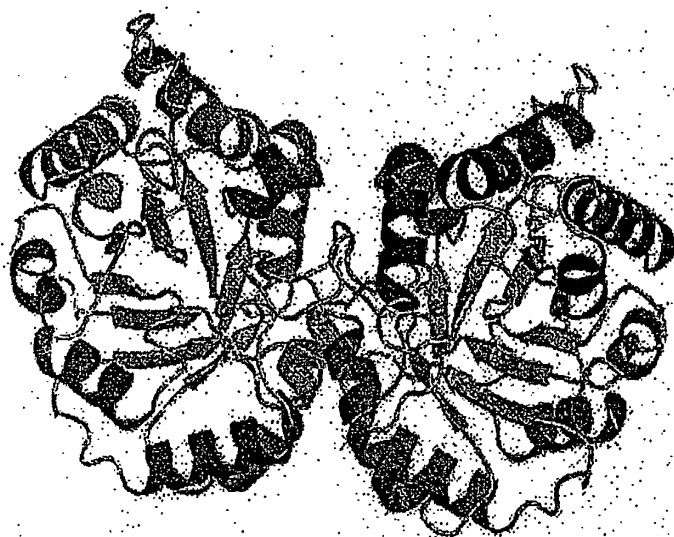


FIGURE 13



72/76

FIGURE 14

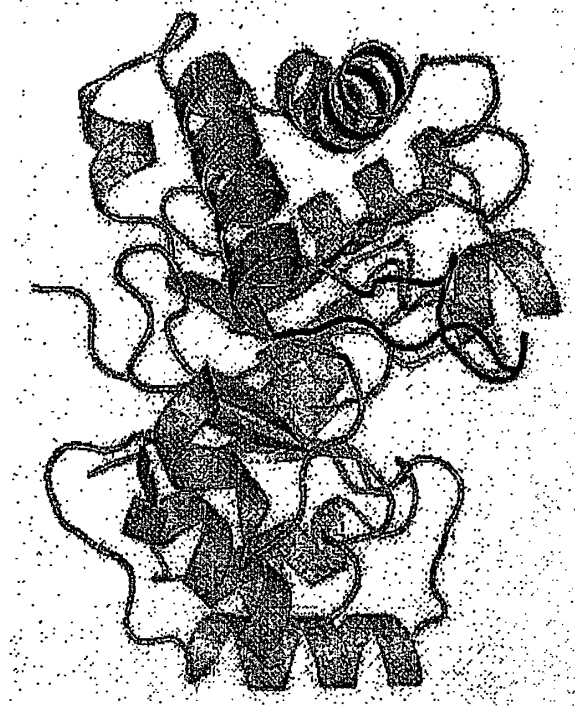
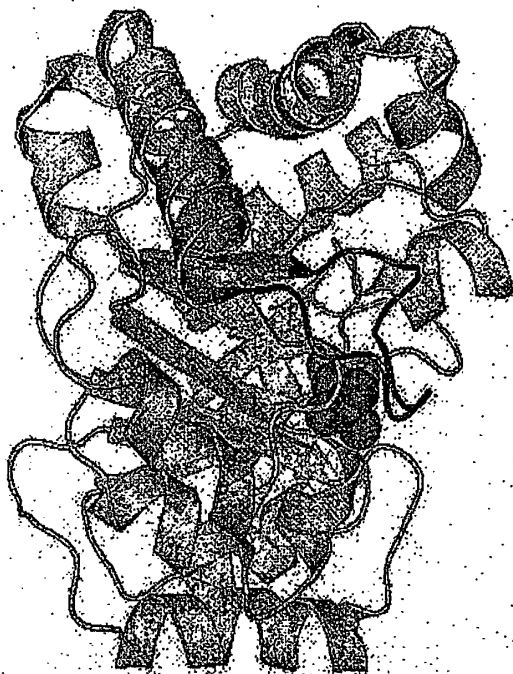


FIGURE 15



74/76

FIGURE 16

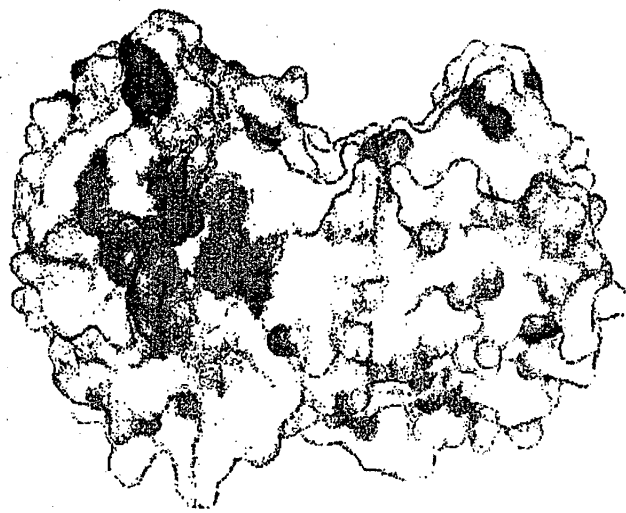


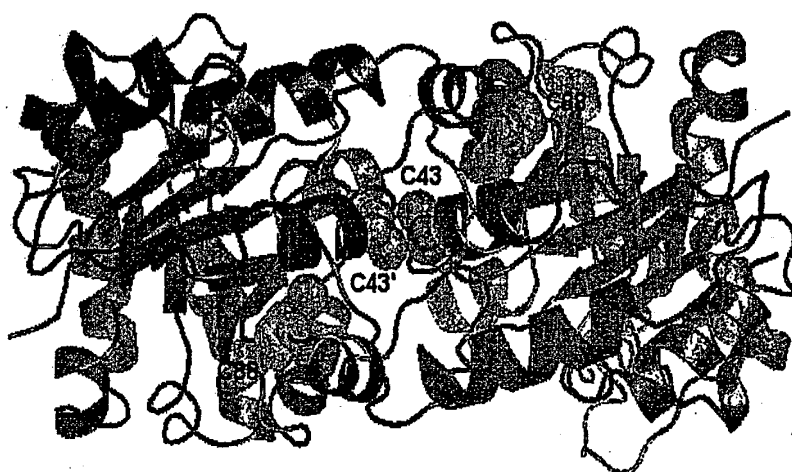
FIGURE 17

TABLE 4: Interface residues of TcTIM, PaTIM, and HsTIM

| Tc TIM | Pa TIM | Hs TIM |
|--------|--------|--------|
| N12 | N9 | N11 |
| K14 | K11 | K13 |
| C15 | M12 | M14 |
| N16 | H13 | N15 |
| G17 | G14 | G16 |
| S18 | T15 | R17 |
| E19 | H16 | K18 |
| S20 | S17 | Q19 |
| T45 | P42 | P44 |
| F46 | C43 | T45 |
| L47 | L44 | A46 |
| H48 | F45 | Y47 |
| I49 | I46 | I48 |
| P50 | S47 | D49 |
| M51 | Q48 | F50 |
| Q66 | Q63 | Q64 |
| N67 | N64 | N65 |
| I69 | V67 | V67 |
| S72 | Q71 | N71 |
| G73 | G72 | G72 |
| A74 | A73 | A73 |
| F75 | L74 | F74 |
| T76 | T75 | T75 |
| G77 | G76 | G76 |
| E78 | E77 | E77 |
| V79 | T78 | I78 |
| S80 | A79 | S79 |
| I83 | Q82 | M82 |
| L84 | L83 | I83 |
| D86 | D85 | D85 |
| Y87 | V86 | C86 |
| I89 | C88 | A88 |
| V93 | L92 | V92 |
| H96 | H95 | H95 |
| E98 | E97 | E97 |
| R99 | R98 | R98 |
| Y102 | I101 | V101 |
| Y103 | L102 | F102 |
| I173 | I172 | I170 |
| G174 | G173 | G171 |

76/76

FIGURE 18



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
3 April 2003 (03.04.2003)

PCT

(10) International Publication Number
WO 03/027274 A3

(51) International Patent Classification: C12N 9/90,
C07K 14/21, 19/00, 17/02, C12Q 1/533, C12N 1/21 //
C07K 103/00, 101/00

(21) International Application Number: PCT/CA02/01453

(22) International Filing Date:
25 September 2002 (25.09.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/324,739 25 September 2001 (25.09.2001) US
60/376,917 1 May 2002 (01.05.2002) US

(71) Applicant (for all designated States except US):
AFFINIUM PHARMACEUTICALS, INC. [CA/CA];
100 University Avenue, 10th Floor, South Tower, Toronto,
Ontario M5J 1V6 (CA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): EDWARDS, Aled
[CA/CA]; 21 Sutherland Drive, Toronto, Ontario M4G
1H1 (CA). DHARAMSI, Akil [CA/CA]; 29 Moresby
Street, Richmond Hill, Ontario L4B 4K9 (CA). VEDADI,
Masoud [CA/CA]; 178 Northwood Drive, Toronto, On-
tario M2M 2K3 (CA). ALAM, Muhammad, Zahoor
[PK/CA]; #11-440 Malaga Road, Oshawa, Ontario L1J
1N8 (CA). AWREY, Donald [CA/CA]; 2211 Stir Cres-
cent, Mississauga, Ontario L4Y 3V2 (CA). BEATTIE,

Bryan [CA/CA]; 2224 Vista Oak Road, Oakville, Ontario
L6M 3L7 (CA). HOUSTON, Simon [GB/CA]; 100
Vaughan Road, Apt. 29, Toronto, Ontario M6C 2M1
(CA). KIMBER, Matthew [CA/CA]; 306 Sumach Street,
Apt. 7, Toronto, Ontario M5A 3K2 (CA). LAM, Robert
[CA/CA]; 85 Thorncliffe Park Drive, Apt. 1505, Toronto,
Ontario M4H 1L6 (CA). NETHERY, Kathleen [CA/CA];
1185 Lansdowne Avenue, Lower Unit, Toronto, Ontario
M6H 3Z7 (CA). LI, Qin [CA/CA]; 3455 Rue Durocher,
Apt. 314, Montreal, Québec H2X 2C9 (CA). NECAKOV,
Sasha [CA/CA]; 530 Brunswick Avenue, Toronto, On-
tario M5R 2Z5 (CA). VALLEE, Francois [FR/CA]; 490
Eglinton Avenue East, Apt. 405, Toronto, Ontario M4P
1M4 (CA).

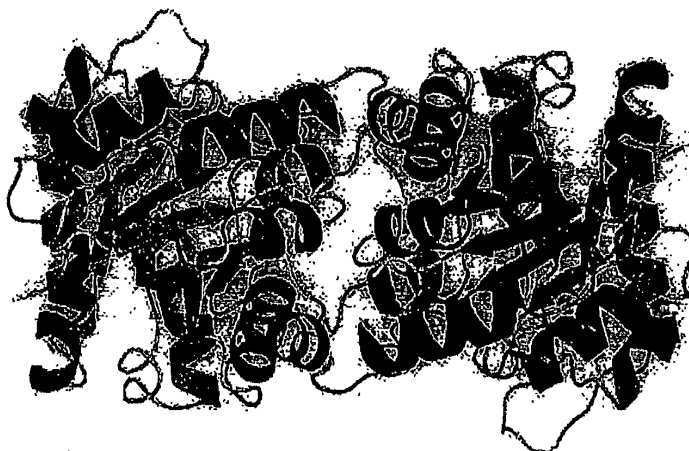
(74) Agent: BERESKIN & PARR; 40 King Street West, 40th
Floor, Toronto, Ontario M5H 3Y2 (CA).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK,

[Continued on next page]

(54) Title: PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA



(57) Abstract: The present invention relates to drug targets for pathogenic bacteria. Accordingly, the invention provides purified protein comprising the amino acid sequence set forth in SEQ ID NO: 4. The invention also provides biochemical and biophysical characteristics of the polypeptides of the invention.

WO 03/027274 A3



TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:

2 October 2003

INTERNATIONAL SEARCH REPORT

International Application No

PCT/CA 02/01453

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N9/90 C07K14/21 C07K19/00 C07K17/02 C12Q1/533
 C12N1/21 //C07K103:00,C07K101:00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS, EPO-Internal, CHEM ABS Data, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Y | <p>STOVER C K ET AL: "COMPLETE GENOME SEQUENCE OF PSEUDOMONAS AERUGINOSA PA01, AN OPPORTUNISTIC PATHOGEN" NATURE, MACMILLAN JOURNALS LTD. LONDON, GB, vol. 406, no. 6799, 31 August 2000 (2000-08-31), pages 959-964, XP000996980 ISSN: 0028-0836</p> <p>The full genomic sequence was disclosed and could be retrieved at www.pseudomonas.com; the Swall entry ID TPIS_PSEAE shows 100% identity with SEQ ID N°2 and 99.2% identity with claimed SEQ ID N°4</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/--</p> | 1-41, 56-69, 71 |

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

Z document member of the same patent family

Date of the actual completion of the international search

7 August 2003

Date of mailing of the international search report

21/08/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Vix, 0

INTERNATIONAL SEARCH REPORT

International Application No

PCT/CA 02/01453

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|
| Y | <p>WILLIAMS JOHN C ET AL: "Structural and mutagenesis studies of leishmania triosephosphate isomerase: A point mutation can convert a mesophilic enzyme into a superstable enzyme without losing catalytic power." PROTEIN ENGINEERING, vol. 12, no. 3, March 1999 (1999-03), pages 243-250, XP002250090 ISSN: 0269-2139 page 244-246</p> | 1-41, 56-69,71 |
| Y | <p>VELANKER SAMEER S ET AL: "Triosephosphate isomerase from Plasmodium falciparum: The crystal structure provides insights into antimalarial drug design." STRUCTURE (LONDON), vol. 5, no. 6, 1997, pages 751-761, XP009015082 ISSN: 0969-2126 see Material and methods and pages 751-752, 758</p> | 1-41, 56-69,71 |
| Y | <p>WIERENGA R K ET AL: "COMPARISON OF THE REFINED CRYSTAL STRUCTURES OF LIGANDED AND UNLIGANDED CHICKEN YEAST AND TRYPA NOSOMAL TRIOSEPHOSPHATE ISOMERASE" JOURNAL OF MOLECULAR BIOLOGY, vol. 224, no. 4, 1992, pages 1115-1126, XP009015108 ISSN: 0022-2836 the whole document</p> | 1-41, 56-69,71 |
| Y | <p>NIENABER VICKI L ET AL: "Discovering novel ligands for macromolecules using X-ray crystallographic screening." NATURE BIOTECHNOLOGY, vol. 18, no. 10, October 2000 (2000-10), pages 1105-1108, XP002250092 ISSN: 1087-0156 the whole document</p> | 40,41, 56-69,71 |
| Y | <p>GSCHWEND D A ET AL: "MOLECULAR DOCKING TOWARDS DRUG DISCOVERY" JOURNAL OF MOLECULAR RECOGNITION, HEYDEN & SON LTD., LONDON, GB, vol. 9, 1996, pages 175-186, XP000882526 ISSN: 0952-3499 the whole document</p> | 19,25, 28-32, 36-41, 56-69,71 |

-/--

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 02/01453

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| P,X | WO 01 70955 A (YAMAMOTO ROBERT T ;OHLSEN KARI L (US); WALL DANIEL (US); XU H HOWA) 27 September 2001 (2001-09-27) Protein in example 3, SEQ ID N°12053 shows 100% identity with SEQ ID N°2 in 251 aa overlap. | 1-41, 56-69,71 |
| P,Y | ----- JOUBERT F ET AL: "Structure-based inhibitor screening: A family of sulfonated dye inhibitors for malaria parasite triosephosphate isomerase." PROTEINS, vol. 45, no. 2, 1 November 2001 (2001-11-01), pages 136-143, XP009015107 ISSN: 0887-3585 especially page 137-139 and Fig. 7 | 1-41, 56-69,71 |
| A | ----- FLEISCHMANN WOLFGANG ET AL: "A novel method for automatic functional annotation of proteins." BIOINFORMATICS (OXFORD), vol. 15, no. 3, March 1999 (1999-03), pages 228-233, XP002250093 ISSN: 1367-4803 the whole document | 1-12 |
| P,A | ----- BRENNER STEVEN E: "A tour of structural genomics." NATURE REVIEWS GENETICS, vol. 2, no. 10, October 2001 (2001-10), pages 801-809, XP002250126 ISSN: 1471-0056 the whole document ----- | 1-41, 56-69,71 |

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA 02/01453

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 42-55, 70, 72-75
because they relate to subject matter not required to be searched by this Authority, namely:
Rule 39.1(v) PCT - Presentation of information
2. ☒ Claims Nos.: 42-55, 70, 72-75
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 42-55, 70, 72-75

Concerning claims 42-55, 70, 72-75 applicant's attention is drawn to Rule 67(v) PCT. The subject-matter of claims 42-55, 70, 72-75 refers to the presentation of structure data (a structure model of the P. aeruginosa triosephosphate isomerase), scalable three-dimensional configuration of points, and is not regarded as patentable invention within the meaning of Rule 67 (v) PCT since it relates to a presentation of information (protein model structure coordinates or configuration of points characterised by coordinates) as a coordinate listings, or information stored on a computer or computer readable media. Thus, the above mentioned claims will not be searched.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/CA 02/01453

| Patent document cited in search report | | Publication date | Patent family member(s) | Publication date |
|-------------------------------------------|---|---------------------|----------------------------|---------------------|
| WO 0170955 | A | 27-09-2001 | AU 4934501 A | 03-10-2001 |
| | | | CA 2404260 A1 | 27-09-2001 |
| | | | EP 1268774 A2 | 02-01-2003 |
| | | | WO 0170955 A2 | 27-09-2001 |
| | | | US 2002061569 A1 | 23-05-2002 |

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.